

# SEARCH REQUEST FORM

49403

Requestor's Name: \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Date: \_\_\_\_\_ Phone: \_\_\_\_\_ Art Unit: \_\_\_\_\_

## Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

## STAFF USE ONLY

Date completed: <u>8/21</u>	Search Site	Vendors
Searcher: <u>P. Schreiber 368-4292</u>	<input type="checkbox"/> STIC	<input type="checkbox"/> IG
Terminal time: <u>8</u>	<input checked="" type="checkbox"/> CM-1 12 E 18	<input type="checkbox"/> STN
Elapsed time: <u>5</u>	<input type="checkbox"/> Pre-S	<input type="checkbox"/> Dialog
CPU time: _____	Type of Search	<input type="checkbox"/> APS
Total time: _____	<input type="checkbox"/> N.A. Sequence	<input type="checkbox"/> Geninfo
Number of Searches: _____	<input type="checkbox"/> A.A. Sequence	<input type="checkbox"/> SDC
Number of Databases: <u>9</u>	<input type="checkbox"/> Structure	<input type="checkbox"/> DARC/Questel
	<input type="checkbox"/> Bibliographic	<input checked="" type="checkbox"/> Other <i>Compu</i>

**THIS PAGE IS BLANK**



49403

Schreiber, David

From: Zhou, Shubo (AU1631)  
Sent: Wednesday, August 15, 2001 12:08 PM  
To: Schreiber, David  
Subject: seq search for 09/598,042

Hi David,

Defective CRF

Another seq search where I'd like to search 09/552,317 to see if the sequence is in that case or not.

Thanks,

Joe

Shubo "Joe" Zhou, Ph.D.  
Patent Examiner  
(703)-605-1158, CM1/12B03  
AU 1631, US PTO

## Search Request

\*\*\*\*\*

*Requester's full name:* Shubo "Joe" Zhou *Examiner #:* 78282

*Art Unit:* 1631 *Phone #:* 703-605-1158 *Mailbox #:* 12D01/CM1

*Results format:* pape *Room #:* 12B03

\*\*\*\*\*

**Serial #: 09/598,042**

**Please search:**

**Nucleic acid  
and  
Protein**

databases for

**SEQ ID NO: 2**

Including:

**1. default search**

**THIS PAGE IS BLANK**

49403

**2. Interference search but only search 09/552,317**

**Please provide 45 alignments for the default search.**

**THIS PAGE IS BLANK**



AX080801	Sequence
AX099939	Mus musc
AC019103	Homo sap
AC070329	Sequence
AC013293	Homo sap
AC080007	Homo sap
AC000380	Homo sap
AC069450	Homo sap
AF110520	Mus musc
AF100956	Mus musc
AC070304	Homo sap
AC027728	Homo sap
AC023162	Homo sap
AL031709	Homo sap
AF111170	Homo sap
AL513355	Homo sap
U03840	Human fibro
X57200	Human fibro
A27111	Tyrosine K1
AC005737	Homo sap
M24637	Chicken tyr
AF245114	Homo sap
AX047615	Sequence
AX047599	Sequence
AX047599	Sequence
AX047600	Sequence
AX047601	Sequence
AX047602	Sequence
AX047603	Sequence
AX047604	Sequence
AX047605	Sequence
AX047606	Sequence
AX047607	Sequence
AX047608	Sequence
AX047609	Sequence
AX047610	Sequence
AX047611	Sequence

[illegible]

[illegible]

Db	1741	CGCCACACACGGTCTCCTTGAAGCCACACAGTACACACACGACACATGACACAGATATGCGC	1800
OY	1889	ccttggacaacaagaataagcttgcaccaatgaacagcaacagacagagatagtcgaagaat	1948
Db	1801	CTTGCGACACACAGATATAGTGTGCCCCAATATGCACGCACACGCAACAGAGACATGCGACAACT	1860
OY	1949	acaaagacatgtctcttgaacatacacacgcgacacccatctgcgcagatgttgtctctgcac	2008
Db	1861	ACAAGGACATGTCCTCTGAACATACACAGCAGCACACCCTATGCGCATGTGCTGCTGGAC	1920
OY	2009	acacacacacacacagatatagtctgtctgcagcacaacgcgtgagatatagtataccggac	2068
Db	1921	ACAACACACACACAGGATATGCTGTGTGGACGCACACAGTCAATATGTATATCCGAC	1980
OY	2069	acacacgtgacacagatatagtctgtctgcagacacagataatgtctgttcgttgcacacacatg	2128
Db	1981	ACACACGTGCACAGATATGCTGTCTGTGACACACACAGATATATGCTTGTGACACACACTG	2040
OY	2129	cacgcatattgtccctggagacacacacacacacacagcttgcgcagagatatagtctgtgcagc	2188
Db	2041	CACGATATTTGCTGTGGACACACACACACACACGCGGTGCACACATATGTGTGTGGACAG	2100
OY	2189	cacacacatgcagatatagtctgtctgcagacacacacttcagacacacgttgcagagcgac	2248
Db	2101	CACACACATGCAGATATGCTGCTGTGACACACACTTCCAGACACACGTGCACAGCGCAG	2160
OY	2249	ataatgtgcctggagacacacgcgagatatagtctgtctagtaacacacacgcgagacatgtg	2308
Db	2161	ATATGTCTCTGGACACACCCACATATGCTGTGTACTACACACACACGCAACGCAACTGTG	2220
OY	2309	tcgcgacacacacacacgcgatgcagagatatagtcttccgcgacacacacgcgcgcagatata	2368
Db	2221	TCCGGACACACACGCGATCTACAGATATGTGCTTCCGGACACACACAGCAGCAGATAT	2280
OY	2369	gcgtgcctggagacacacacagataatatagtctgtcctcaacactcacacgcgtgcagatattgc	2428
Db	2281	GCTCCTCGGACACACACACAGATATATGCTGCTCAACACTCCACACACGTCGACAGATATGC	2340
OY	2429	ctggacacacacacatgtgcagagataatgtgctgtctgcagacatgcagacacgcgtgcagatagt	2488
Db	2341	CTGGACACACACATGTACACAGATATGTCTGTGACATGACACACACGTCGACAGATATGCT	2400
OY	2489	gtccgcgatatcacacgcgcacacacatgcagagatatagtctgtctgtgcgcacacacttcgcgcga	2548
Db	2401	GTCGGATACACACACGACGACACACACATGTGCAGATATGCTGTGCGCACACACTTCCGGACA	2460
OY	2549	cacatgcacacacacagctgcagatatagtctgtcctgcagacacgcgcagactgcgtcttttg	2608
Db	2461	CACATGCACACACACGTCGAGATATGCTGCTGTGACACACGACAGACTGACGCTTTTGG	2520
OY	2609	gaagagtgtgcgtgaagcgcgcagctacgtgtgcgtggaagtcacatagtgtatgttgagact	2668
Db	2521	GAGGAGTGTGCGGTAAACCTGTGACGTAGTGTGCGGTGAGAGGCTCATATGTTGATGAGGACT	2580
OY	2669	ttccctgtccacgcgtcaactcccccacacttgcgcgcctctgttccgcgccttaagtcgccg	2728
Db	2581	TTCCCTGTCTCACCGTTCACCTCCCAACTGTGCGCGCTCTTCCCGCTCATAGTCCCGC	2640
OY	2729	cctccatcccccgcctgtgccctgtgcctgtgcgctattttgcacacgtgccttgggtg	2788
Db	2641	CTCTCATCCCGCCTCTGTGCTCCCTCGGACCTTGGCGGCTATTTTGTGCACCTGCTTGGGTTG	2700
OY	2789	cccaagagctccctcactgcctgttgggtgttgggtgttgggggagacagccccaagccttga	2848
Db	2701	CCCAAGAGTCCCTACTGCTGTGGGCTGTGGGGGTGGGGGCACAGCAAGCCCCAAGCTTGAGA	2760
OY	2849	gactgcgagcccatgtcagctgtgcctacccacatgcatttcccccctgcacacagagaag	2908
Db	2761	GGCTGAGCCCATGTGCTAAGTGGCTCATCCCCACTGTCTTCCCTTGACACAGAGAAG	2820
OY	2909	ggccttgcattataatlaagaatagaaataataataaagatgcgaaggaagctg	2968





[illegible]

Db	2646	CAACACTCACACACCTGGAGATATTTGGCTGGACACACACATGTGACAGAGATATGCTGCT	2705
QY	2462	ggagacatgacacacagctgacagatatgtctccggaatacacacagcaacacatgcaagata	2521
Db	2706	GGACATATGCACACACGTGCGATATCTGTCCGGATACACACGACACACATGACAGATA	2765
QY	2522	tgctcctgtggacacacattccgagacacacatgacacacaggtgacagatatgtctgcctg	2581
Db	2766	TGCTGCTGTGGGACACACATCTCCGGACACACATGACACACAGGTGACAGATATGCTGCTG	2825
QY	2582	gacacacgacagctacgtgtcctcttgggaagggtgtgcgtgaagcctgcagtaagttgtgc	2641
Db	2826	GACACACCCAGACTACAGCTGCTTTTGGGAGGGGTGTGCCGTGAAGCTCAGTACCTGTGC	2885
QY	2642	cgtagagctcatagttgtatgaaggacttccctgtctcaacgttaactccccaaactgcgc	2701
Db	2886	CGTAGAGCTCATAGTTGATGAAGGACTTCCCTGCTCACCGCTACTCCCCCACTCTGC	2945
QY	2702	ccgctctgtcccgccctcagtcctccgcctccatcccgctctgtccctctgcctgtgc	2761
Db	2946	CGGCTGTGTGCCCGGCTACAGTCCCGGCTGCATCCCGGCTGTGCTGCCGCTTGCC	3005
QY	2762	ggctatttttgcacactgcttgggtgtcccaaggatccctactgtcgttggcttgggtt	2821
Db	3006	GGCTATTTTTTGGCACCTGCTGTGGGTGCCACAGACTCCCTACTGCTGTGGGCTGGGCTT	3065
QY	2822	gggggacagacagaccccaagccttggaaggcttggagcccatgctagttgactcatcccaac	2881
Db	3066	GGGGGACACAGCACGCCCAAGCTGTGAAGGCTGGAGCCCATGCTAGTGGCTATCCCGAG	3125
QY	2882	tgcatctccccccttacacagagaaggggccttggatattatattaaagaatgaagata	2941
Db	3126	TGCATTTCTCCCTGACACAGAGAGAGGGGCTTGCTATTATTTAAGAAATGAAGATA	3185
QY	2942	atattaataatgataatgaagaagactggtgttggagaagcttggcttccctggggcccg	3001
Db	3186	ATATTAAATATATATGGAAGAGAACTGGGTGTCAGAGGACTGTGGTCTCTCTGGGGCCCG	3245
QY	3002	ggagccgcctgtgtctttaaagccatgctatgagcacacaccccgatcagaagccagaacacac	3061
Db	3246	GGAGCCGCTGTGCTTTTACGCCATCTGTGACACACACCCGCTCCAGGCGACACACACC	3305
QY	3062	ccccacccacactgtgtgtgtgccccagagatctgttaatttatgtagaagttgagctga	3121
Db	3306	CCCCACCCACAGTGTGTGTGTGGCCCAAGTCTCTTAATTTATATGTAAGTTTGAAGCTGA	3365
QY	3122	agccccgatatttaattatttgtttaaacaatga	3158
Db	3366	AGCCCCGATATTAATTAATTAATTTTGTAAACACAAAA	3402
RESULT 3			
AF312678	3397	bp	02-MAR-2001
LOCUS	AF312678	3397 bp	02-MAR-2001
DEFINITION	Homo sapiens FGF homologous factor receptor (FHR)	complete cds.	complete
ACCESSION	AF312678		
VERSION	AF312678.1	GI:13183617	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 3397)		
	Aggarwal,S., Xie,M.-H., Foster,J., Frantz,G., Stinson,J.,		
	Corpusz,R.T., Simmons,L., Hillan,K., Yansura,D.G., Vandlen,R.L.,		
	Goddard,A.D. and Gurney,A.L.		
TITLE	FHR, a novel fibroblast growth factor receptor that uniquely binds the fibroblast growth factor homologous factors		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 3397)		
AUTHORS	Aggarwal,S., Xie,M.-H., Foster,J., Frantz,G., Stinson,J.,		



[illegible]

OY	2642	cgttagagcctcaatagtgtgatgaaggaccttccctgctccaacgctaactccccaaactctgc	2701
Db	2865	CGTGGAGGCTCATAGTTGTGAAGAAGACTTTTCCTGTGCCACGCTGCATCCCCCAACTCTGC	2924
OY	2702	cggccctctctcccgcgcctcaagtcaccgcgcctcccatcccccgcctctctccccttgacctgc	2761
Db	2925	CGGCTCTGTCCCCGCTCAGTCAGTCCCGCTCATCCCGCTCTGTGCCCTTGCTGGCTTGGC	2984
OY	2762	ggcatttttgtccacactgaccttggattgcaccaagaatgccctaactctgcttggcgctggatt	2821
Db	2985	GCGTAATTTTGGCACACTGCTTGGGATGCCAGAGAGGCCCTCACTGCTGAGGCTGGGCTGGATT	3044
OY	2822	gggggacaagcacgcccccaagccttagaaggctgtagaacccatactgctagtgtctatcccac	2881
Db	3045	GGGGGACAAGCAGCCCCAACGCTGAGAGGCTTGAGAGCCATGCTAAGTGCTCATCCCGAG	3104
OY	2882	tgcattctccccttacacagagaagaaggccttgatttatcatatcaaagaata	2941
Db	3105	TGCATTCTCCCTTGACACAGAGAGGGCTTGGATTATTTAATTTAAGAAATGAAGATA	3164
OY	2942	atatataaatatgatygaaaagaagactygglttgcaggagacttgcctctcctcggggccgc	3001
Db	3165	AATATTAATTAATGATGGAAGAGAGACTGGGTTCAGAGGACTGAGTCTCCTGGGGCCCG	3224
OY	3002	ggaccgcgcgtgtcttccagccatgctgatgacacacaccccgltccaggccagacacacc	3061
Db	3225	GGACCCTGCTGTCTTTTCAGCCATGCTGATGACACACCCCTCCAGGCGACACACACAC	3284
OY	3062	ggcaccaccacgtctgtgtgagcccaagatctctgaatttatcttaagqtltaggttgcctga	3121
Db	3285	CCCCACCCCACTGTCGTGTGGCCCCAGATCTCTGTAATTTATGATGAGTTTAGACTGA	3344
OY	3122	agcccgatlatltaattatcttgtltaacaatgana	3158
Db	3345	AGCCCCGTAATTTAATTTATTTTGTAAACACAAA	3381
RESULT 4			
AKO22597			
LOCUS	AKO22597	1999 bp mRNA	PRI 29-SEP-2000
DEFINITION	Homo sapiens CDNA FLJ12535 fls, clone NT2RM4000251.		
ACCESSION	AKO22597		
VERSION	AKO22597.1	GI:10434078	
KEYWORDS	oligo capping; fls (full insert sequence).		
SOURCE	Homo sapiens testiscarcinoma cell line:NT2 CDNA to mRNA, clone.lib.NF2RM4 Clone:NT2RM4000251.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (sites)		
AUTHORS	Isozaki,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Wadatsuma,M., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiyuchi,S., Watanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nageharu,K., Mashio,Y., Nimomiya,K. and Iwayanagi,T.		
TITLE	NPO human cDNA sequencing project		
JOURNAL	Unpublished (2000)		
REFERENCE	2 (bases 1 to 1999)		
AUTHORS	Isozaki,T. and Otsuki,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-AUG-2000) to the DDBJ/EMBL/GenBank databases. Takao Isozaki, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genominfo@ri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952) NEBO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan: cDNA full insert sequencing: Research Association for Biotechnology: cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.		

FEATURES		Location/Qualifiers
Source		1. 1999 /organism="Homo sapiens" /db_xref="taxon:9606" /cell_line="NT2" /cell_type="teratocarcinoma" /clone="NT2RM4000251" /clone_11b="NT2RM4" /note="Cloning vector: pME18SFLJ-mRNA from uninduced NT2 neuronal precursor cells."
BASE COUNT	463 a 666 c 509 g 361 t	
ORIGIN		
Query Match	58.3%;	Score 1856.2; DB 89; Length 1999;
Best Local Similarity	93.8%;	Pred. No. 0;
Matches 1996; Conservative	0;	Mismatches 3; Indels 128; Gaps 1
QY 1049	tgcacagaccacaaacccgacgagggccacacctgtgacctctcgtctctcgcgacctaagctgc	1108
DB 1	tgcacagaccacaaacccgacgagggccacacctgtgacctctcgtctctcgtcgcgacctaagctgc	60
QY 1109	cgttgcccgctggtcatalctggtcatccacccgctgtcttcatctctgggcaacctgtctcc	1168
DB 61	cgttgcccgctggtcatalctggtcatccacccgctgtcttcatctctgggcaacctgtctcc	120
QY 1169	tgtgtgctttgcagggcccgaaagaacgctgtacccccgcgcctctgccccctccctctcctg	1228
DB 121	tgtgtgctttgtgcacggccccaagaaacccctgtcacccccgcgcctctgccccctccctctcctg	180
QY 1229	ggcacccgcccgcggyggaacggtcccgagccgcgacgagagagaaagaaacctccctgttg	1288
DB 181	ggcacccgcccgcggyggaacggtcccgagccgcgacgagagagagaaagaaacctccctgttg	240
QY 1289	ccgacctcagcgctgtgacctgtgtgtgggtctgtgtgagagcaatggtctccgcaagccc	1348
DB 241	ccgacctcagcgctgtgacctgtgtgtgggtctgtgtgagagcaatggtctccgcaagccc	269
QY 1349	cccagcacttactgtggcccagggccaggttgtgacctaaagtgtaccccccaactctaca	1408
DB 270	-----	269
QY 1409	cagacatccaacacacacacacactgtatctgcygcgcgctgtgtgagagagatgtggtctcc	1468
DB 270	-----	292
QY 1469	ggcagccccccagacttactctggccacagcccaagctgtctgcgcctaagtgtgaccccaa	1528
DB 293	ggcagccccccagacttactctggccacagcccaagctgtctgcgcctaagtgtgaccccaa	352
QY 1529	actctacacagacatccaacacacacacacacacacactctcaacacacatccaactgtgga	1588
DB 353	actctacacagacatccaacacacacacacacacacactctcaacacacatccaactgtgga	412
QY 1589	gggcaaggtccacacagacacatccactctcaagtgtgtaacggcaacctgtctgcaagagggc	1648
DB 413	gggcaaggtccacacagacacatccactctcaagtgtgtaacggcaacctgtgtgcaagagggc	472
QY 1649	acgggggggcccggccagacagacagcagactctggagagatgtgagagacggactcgaacgaag	1708
DB 473	acgggggggcccggccagacagacagcagactctggagagatgtgagagacggactcgaacgaag	532
QY 1709	cagggggaacctgtgcggaggaggaatgtgcccagcccccaagcgaactctgtgtgtggcata	1768
DB 533	cagggggaacctgtgcggaggaggaatgtgcccagcccccaagcgaactctgtgtgtggcata	592
QY 1769	ggcccttgacacacacacacacacacacacactcctgtgactgtgcatgtgtgacacacatg	1828
DB 593	ggcccttgacacacacacacacacacacacactcctgtgactgtgcatgtgtgacacacatg	652
QY 1829	cgtgcacacgtgtctccctgaaagcacaactgtgcacacaogcacatgtgacagatatgcg	1888
DB 653	cgtgcacacgtgtctccctgaaagcacaactgtgcacacaogcacatgtgacagatatgcg	712

QY	1889	cctggtgacacagataagctgtcccaatgacgcacgcacagagaacatgtccagaacct	1948
Db	713	cctgtggcacacagatTAAGCTGGCCAAATGCAAGCACACGACACAGACATGACCAAGACAT	772
QY	1949	acaaagacatgctgtcgtgaaacatacaacgcacaccacatgacagatgtgtctccgagac	2008
Db	773	ACAAAGACATAGCTGGCCGTAACATACACGACACACCCATGGCGCAGATGTCTGCTGGAC	832
QY	2009	acacacacacacacagatatgtctgtctgtgacgcacacacgtgtcagatatgtatccggac	2068
Db	893	ACACACACACACACGGATATAGCTGTGGACGCACACACGTGCAGATATGTATCCGGAC	892
QY	2069	acacacgtgcacagatatgtctgtcgtggaacacagatatgtctccttggacacacatg	2128
Db	893	ACACACGTGCACAGATATAGCTGTGGACACACATATATCTGCTGTGACACACATG	952
QY	2129	caagcatattgtccctgagacacacacacacacacgtgtgcacagatatgtcttctgacacg	2188
Db	953	CACGATATTGCTCGGACACACACACACACAGTGTGCACAGATATGCTGTGAGACAG	1012
QY	2189	cacacacatgcagatatgtctgtcgtgacacacacattccagacacacgtgtcacaggcgag	2248
Db	1013	CACACACATGCAGATATAGCTGTGGACACACACTTCCAGACACACGTGCACAGCGCCAG	1072
QY	2249	ataatgtgtcctggagacacacacgcggaatgtgtctatgtacaacacacacgcagacatgctg	2308
Db	1073	ATATGTGCTGTGGACACACGCAGATATGCTGTCTGTAGTCAACACACACAGACAGATGCTG	1132
QY	2309	tccggacacacacacgcgatgcagagatalgtctgtctcggagacacacacgcacgcagatat	2368
Db	1133	TCCGGACACACACACGCATGCACAGATATAGCTGTCCGACACACACACGCGAGATAT	1192
QY	2369	gtctgctgtgagacacacacagataatagtcgtccttaaacctcaacacacgtgtgagatalgtc	2428
Db	1193	GCTGCTGTGACACACACAGATATAGTGTGCTCTCAACACTCAACACGTGCAGATATTGC	1252
QY	2429	ctgagacacacacatgtgcagagatalgtctgtctgtgacatgtcacacacgtgtcagatalgtc	2488
Db	1253	CTGGACACACACATGTGCACAGATATAGTGTGTGGACATGACACACGTGCAGATATGCT	1312
QY	2489	gtccgagatacacacgcacgcacacacacatgcagatatgtctgtcgtggagacacattccggaca	2548
Db	1313	GTCCGAGTATACACGCGACGACACATGCAGATATGCTGTGGACACACTTCCGACACA	1372
QY	2549	cacacgtcacacacacggtgcacagatalgtcgtccgtgacacacgcagacatgcgttcttgg	2608
Db	1373	CACATGCACACACAGGTGCAGATATGCTGTGGACACACGACACTGACGTGCTTTGG	1432
QY	2609	gagggatgtgcgtgaaagcgtgcagtaacgtgtgcgtgaggtccataagttgattggagac	2668
Db	1433	GAGGGTGTGCCGTAAAGCTGCAGATAGTGTGCCGTGAAGGCTCAATGATGATGAGGACT	1492
QY	2669	ttcctcgttcacacgtaactcccccaactctgcgcgcctctgttcccgcgcctcaatccccg	2728
Db	1493	TTCCCTGCTCCACCGTCACTCCCAACTGTGCCCGCCTGTTCGCCGCTCAAGTCCCG	1552
QY	2729	ctctcaatcccgctctgtccctgtgcgtgtgcgtatattgtccacatgcgtcttggtg	2788
Db	1553	CTTCATCCCGCGCTGTCTGCCCTGTGGCCTTGGCGCTTATTTTGGACCTGCTTGGGTG	1612
QY	2789	cccagagatccctactatgctgttggcgtggggtttgggggacacacagccccaagccttga	2848
Db	1613	CCCAGAGTCCCTTAATGCTTGTGGGCTGGGGTTGGGGGCACACAGCCCAAGCTTGAGA	1672
QY	2849	ggctggaagcccatgtcagtggtgcacatccccacatgtatcttccctctgacacagagaag	2908
Db	1673	GGTGGAGCCCATGTGCTATGTGGCTCATCCCACTGCATTCTCCCTGCACACAGAGAG	1732
QY	2909	ggccttgtattatattaaagaatgaatataatataatgatatgaaaggaagatg	2968
Db	1733	GGCCTGTGATTATATTAAAGTAATATATATTATATGATGTGAAGAAAGCTG	1792

OY 2969 gtttcagagagactgtgtctctctctgagagccgagccgctgtcttcagccatgct 3028  
|||||  
Db 1793 GGTTCAGAGGACTGTGTCTCTCTCTGGGCCCCGGAGCCCGCTGTCTTTAGCCATGCT 1852  
OY 3029 gatgacacacacccgcgtccagagccacacacaccccccactgtcgtgtgtgtccca 3088  
|||||  
Db 1853 GATGACACACACCCGTCGAGCAGACACACCCCGCCACTGTGTGTGTGTGTGTGTGT 1912  
OY 3089 gatctcgtatattttagtagagtttgagctgtagagcccgatatttaattatttgtt 3148  
|||||  
Db 1913 GATCTCGTATATTTATGTAGATTGAGCTGAAGCCCGGTATATTTATTTATTTGTT 1972  
OY 3149 aaacatgaagtcgactcttcctccctc 3175  
|||||  
Db 1973 AAACATGAAGTGCATCTTCTCCCTCC 1999

RESULT 5  
AC019103/c  
LOCUS  
DEFINITION Homo sapiens chromosome 4 clone RP11-460119, WORKING DRAFT  
AC019103  
ACCESSION  
VERSION AC019103.5 GI:9838209  
KEYWORDS HTG: HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 168844)  
Waterston,R.H.  
The sequence of Homo sapiens clone  
Unpublished  
2 (bases 1 to 168844)  
Waterston,R.H.  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
On Aug 17, 2000 this sequence version replaced gi:7232144.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Web site: http://genome.wustl.edu/gsc/index.shtml  
----- Project Information -----  
Center project name: H.NH0460119  
----- Summary Statistics -----  
Sequencing vector: M13; 79%  
Sequencing vector: plasmid; 21%  
Chemistry: Dye-terminator ET; 79% of reads  
Chemistry: Dye-terminator Big Dye; 21% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 152408 bases at least Q40  
Consensus quality: 157997 bases at least Q30  
Consensus quality: 160896 bases at least Q20  
Insert size: 176000; agarose-fp  
Insert size: 167848; sum-of-contigs  
Quality coverage: 3.66 in Q20 bases; agarose-fp  
Quality coverage: 3.91 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 24 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 2254: contig of 2254 bp in length  
\* 2255 2354: gap of unknown length  
\* 2355 3899: contig of 1545 bp in length  
\* 3900 3999: gap of unknown length

\* 4000 6569: contig of 2570 bp in length  
\* 6570 6669: gap of unknown length  
\* 6670 8799: contig of 2130 bp in length  
\* 8800 8899: gap of unknown length  
\* 8900 12532: contig of 3633 bp in length  
\* 12533 12633: gap of unknown length  
\* 12633 15196: contig of 2564 bp in length  
\* 15197 15297: gap of unknown length  
\* 15297 18396: contig of 3100 bp in length  
\* 18397 18496: gap of unknown length  
\* 18497 21828: contig of 3232 bp in length  
\* 21829 26181: contig of 4353 bp in length  
\* 26182 26281: gap of unknown length  
\* 26282 28924: contig of 3643 bp in length  
\* 28925 30024: gap of unknown length  
\* 30025 33964: contig of 3940 bp in length  
\* 33965 34064: gap of unknown length  
\* 34065 38781: contig of 4717 bp in length  
\* 38782 38881: gap of unknown length  
\* 38882 45141: contig of 4260 bp in length  
\* 45142 47388: contig of 4147 bp in length  
\* 47389 47488: gap of unknown length  
\* 47489 53220: contig of 5732 bp in length  
\* 53221 53320: gap of unknown length  
\* 53321 57916: contig of 4596 bp in length  
\* 57917 58016: gap of unknown length  
\* 58017 63793: contig of 5777 bp in length  
\* 63794 63893: gap of unknown length  
\* 63894 67854: contig of 3966 bp in length  
\* 67855 67959: gap of unknown length  
\* 67960 75557: contig of 7598 bp in length  
\* 75558 75657: gap of unknown length  
\* 75658 81779: contig of 6122 bp in length  
\* 81780 81879: gap of unknown length  
\* 81880 95438: contig of 13559 bp in length  
\* 95439 95538: gap of unknown length  
\* 95539 109320: contig of 13782 bp in length  
\* 109321 109420: gap of unknown length  
\* 109421 128519: contig of 19099 bp in length  
\* 128520 128619: gap of unknown length  
\* 128620 168844: contig of 40225 bp in length.

FEATURES  
source  
Location/Qualifiers  
1..168844  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="4"  
/clone="RP11-460119"  
1..2254  
/note="assembly\_name:Contig23"  
2255..3899  
/note="assembly\_name:Contig24"  
4000..6569  
/note="assembly\_name:Contig25"  
6670..8799  
/note="assembly\_name:Contig26"  
8900..12532  
/note="assembly\_name:Contig27"  
12633..15196  
/note="assembly\_name:Contig28"  
15297..18396  
/note="assembly\_name:Contig29"  
18497..21728  
/note="assembly\_name:Contig30"  
21829..26181  
/note="assembly\_name:Contig31"  
26282..28924  
/note="assembly\_name:Contig32"  
30025..33964  
/note="assembly\_name:Contig33"  
34065..38781  
/note="assembly\_name:Contig34"







```

QY 2489 gttccgataacacacgacgacacacatgcatatgtgcttgggaacacattccggaca 2548
    |||||||
Db 665 GTCTGGATACACACGACACACATGATGTGCTGGGACACACATCTCCGGACA 724
QY 2549 cacatgcacacacaggtgcagatatgtctgctgagacacacgagactgacgtctttgg 2608
    |||||||
Db 725 CACATGCACACACAGGTGCAGATATGCTGCTGGACACACGAGACTGACGTCTTTGG 784
QY 2609 gaaggtgtgcgtgaagcgtgcagtaagtcgtgctgaggtcctaagttagtagggagct 2668
    |||||||
Db 785 GAGGGGTGCGCTGGAAGCCTGACGTAGTGTGCTGAGAGCTCATAGTTGATGAGGAGCT 844
QY 2669 ttccctgctcacacgtgcaccccccaactctgcccgcctctgtcccgctcagtcacgg 2728
    |||||||
Db 845 TTCCCTGCTCACACGCTGACCTCCCAACTCTGCCCTCTGTGCCCTCAGTCCCG 904
QY 2729 cctccatcccgacgtctgtccctggtcctgtgcgtatatttgcacactgcttgggtg 2788
    |||||||
Db 905 CCTTCATCCCCGGCTGTGCTCCCTGGGCTTGGGGCTATTTTGGCACCTGCTTGGTG 964
QY 2789 cccaggaagtcctcactgtctgtgtggtgtggtgggagcagcagcccaagcctgaga 2848
    |||||||
Db 965 CCCAGGAGTCCCTACTGCTGTGGGCTGGGGTGAGGACACACGCCCAAGCTGAGA 1024
QY 2849 ggtctgagagccatggtgctagtggtcattcccaactgcatctccctctgacagagaag 2908
    |||||||
Db 1025 GGCTGAGCCCATGCTAGTGTGCTCATCCCATGCTTCTCCCTGACACAGAGAAG 1084
QY 2909 ggccttgatattatatttaagaataataataataatgaatgaagaagaactg 2968
    |||||||
Db 1085 GGCTTGTGATTTATATTATTAAGAAATGAATATTAATTAATGATGAGAGAACTG 1144
QY 2969 ggttgcaggaagcgtgtgtctctctctgtggggccggagccgctggtcttcaacactg 3028
    |||||||
Db 1145 GGTGACGAGGACTGTGTCTCTCTGGGGCCCGGACCCGCTGTCTTCAACCATGCT 1204
QY 3029 gatgacacacacccggtcacaagcagacacacaccccaaccactgctggtgagccca 3088
    |||||||
Db 1205 GATACACACACCCGCTCAGGCCACAGACACACCCCACTGCTGTGAGGGCCCA 1264
QY 3089 gatctctgtaatttatgtatgtagttagtgaagcccgatataattattttgtt 3148
    |||||||
Db 1265 GATCTGTGATTTATGTAGAGTTTGAAGTGAAGCCCTATATTAATTTATTTGTT 1324
QY 3149 aaacatgaagtgatctcttccctcccaaaaaaaaaa 3186
    |||||||
Db 1325 AAACATGAAGTGCATCTTAATAAAAAAAAAAAAAA 1362

```

```

FEATURES
Source
Chonju 560-180, Korea
Location/Qualifiers
1..1565
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene
1..1565
/gene="FGFR5"
33..1547
/gene="FGFR5"
/codon_start=1
/product="fibroblast growth factor receptor 5"
/protein_id="AAK26742.1"
/db_xref="GI:13447749"
/translation="MTPSPLLLLLPILLGAPPPAAANGPPPMADKVPVPROVARLGRVRLQCPVEGDPEPLTMTKDRTHISGWSRRVLPQGLKROVEREDAGVYCKATNGFSLSVNTLVLYLDISPKEISLSSGGDEPDAISOOMARPRETOPSKMRRTYARPGSSVRYKLVGSHRPDIIMMKDOLATREAEAPKKWTLSLKLRLPEDSGKYTCVSNRACINATYVYDIQPRSKSPVLTGHPVNTVYDPPGTTSFQCKVSDVPTIOLKRVETGAEGRHNSITDVGSOKRVLPEDVWRSPGSLNKLILTRARODAGMYICLGANTMGYSRSLFTVLDPRKPGPPVASSSATSLPMPVYVIGTAGVFTLGLMLCOAKPKPCTPAPAPPLPGRHPGPTARHSDKDLPSLAASAGPGVGLCEHGSAPAPQHLGPGVAPGPKLYPKLYTDIHTHTHTSHSHVBEKVHQLHYOC"
BASE COUNT
282 a 560 c 497 g 226 t
ORIGIN
Query Match
Best Local Similarity 99.6%; Score 1333.4; DB 89; Length 1565;
Matches 1337; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 89 ttcgttcgagcaggtgtgcccggctgtggccgacatgtcggtctgagtgccagtggaagg 148
    |||||||
Db 139 TCCACAGGACGAGTGCCCGGCTGGGCCACACGTGCTGCTGAGTCCAGTGGAGGGG 198
QY 149 accgcgcgcgcgtgacatgtgtgaccaaagatggtgcgcacacatccacagcgttgaagcc 208
    |||||||
Db 199 ACCCGCCGCCGTGACATGTGTGACCAAGATGGCCGACCATTCACAGCGGCTGAGACC 258
QY 209 gttccgcgtgtgttcgcgcaggggtctgaagtgaaagcaggtgagcggagagatccgcgc 268
    |||||||
Db 259 GCTTCGCCGTGCTGCGCGCAGGGGCTGAAGGTGAACAGGTGAGCGGAGAGATGCGCGCG 318
QY 269 tgaagtgatgaagcgaacacagcgttcgtagctctgaagctgaactaacaccctgcgc 328
    |||||||
Db 319 TGTACTGTGTCAAGGCGACCAAGGCTTGTGCAAGCTTAAGCTCAACTACACCTCTGTCG 378
QY 329 tcttgatgacattagcccaaggaaagagagccttgggcccgcacagctcctctgggggtc 388
    |||||||
Db 379 TGCTGATACATTAAGCCACAGGAGAGAGAGCTGGGGCCGACAGCTCTCTGGGGGTC 438
QY 389 aagaagaccgcgcagcagcagtggtggcagcagcgcgtctcaacagccctccaagatga 448
    |||||||
Db 439 AAGAGAGACCCCGCCACAGCCAGAGTGCGGACGACCGGCTTCAACAGCCCTCAAGATGA 498
QY 449 gggccgggtgtgtcgaagggccgtgtgtagctcgttggtgctcaagtgcgtggcgaagc 508
    |||||||
Db 499 GGGCGCGGATGATGCACGCGCGGTGAGCTCCGTGCGGCTCAAGTCTGTGGCCAGCG 558
QY 509 ggcacccctggcccagcatcacgttgatgaagagcagcagcgttgaagcggcccaagag 568
    |||||||
Db 559 GGCACCTCGGGCCGACATACAGTGTGAAGAGACAGCAGCGCTTGACGCCGCCAGAGG 618
QY 569 ccgctgaagccaggaagaagaagtgtgacactgaagcctgaagaacctgcggcgaggaga 628
    |||||||
Db 619 CCCTGAGGCCAGGAAGAAGAACTGACCTGAGAAACCTGCGCGCGGAGAGAGA 678
QY 629 gggcgaataacacctgcgcgtgtcgaacgcgcgggcgcatacaagcgcacatacaag 688
    |||||||
Db 679 GCGGCAATACACCTCGCGGTTCGAACCGCGGGCCCATCAAGCGCACCTACAAAG 738
QY 689 tggatgtatcagcagcagaccgttccaagccgtgtcacaagcagcagcccgatgaaca 748
    |||||||
Db 739 TGGATGTGATCCAGCGGACCGGTTCCAAAGCCGCTGCTCACAGGACAGCAGCCCGTGACA 798

```

RESULT 7  
 AF279689 1565 bp mRNA PRI 27-MAR-2001  
 LOCUS AF279689  
 DEFINITION Homo sapiens fibroblast growth factor receptor 5 (FGFR5) mRNA, complete cds.  
 ACCESSION AF279689  
 VERSION AF279689.1 GI:13447748  
 KEYWORDS human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1565)  
 Kim, I., Moon, S., Yu, K., Kim, U. and Koh, G. Y.  
 A novel fibroblast growth factor receptor-5 preferentially expressed in the pancreas(1)  
 JOURNAL Biochim. Biophys. Acta 1518 (1-2), 152-156 (2001)  
 MEDLINE 2 (bases 1 to 1565)  
 B1167383  
 AUTHOR KIM, I., MOON, S. and KOH, G.  
 REFERENCE Direct Submission  
 TITLE Submitted (19-JUN-2000) Cardiac Regeneration Group, National Research Initiatives, Chonbuk University, San 2-20, Keum-Am-Dong,



QY 749 cgaagctgagcttcgaggagagcaccgcttcctcagtgcaaggttcgagcagcagctgaagc 808  
DB 799 CGAGGGGTGAGCTTGGGGGGGACCAAGCTCTTCCAGTGCAGAGTGCAGGACAGCTGAAGC 858  
QY 809 cgggagctccagtgctcctaagaagcgctgagagtagcagcgccgagggccgacacacccacca 868  
DB 859 CGGGATTCACATGGCTGAGAGCGCGGGAGTACGGGCGCGAGGGCGCCACACTCCACCA 918  
QY 869 tcgagtggggggcagagaagtttggtgctgcgcacaggggagcagtggtgctggcgcccg 928  
DB 919 TCGATGTGGGGGGCGGACAGAAATTGTGTGCTGCCACAGGGTGAAGTGTGTGGGGCCGG 978  
QY 929 acggctccctacctaataagctgtcctacaccgctgcgcgcaggaagaatgaggagatgt 988  
DB 979 ACGGCTCTTACCTCAATTAAGCTGCTCATACCCCTGCGCCGACAGAGCATGCGGGCATGT 1038  
QY 989 acatctccttgaggcgcaagaacacatgaggtacagcttcgagagcccttcctcaccgtgc 1048  
DB 1039 ACATCTGCTTGGCGGCCACACACATGGGCTACAGCTTCGCAAGCGCTTCTCAACCGTGC 1098  
QY 1049 tgcagaaccacaacacgcagagcagcactgagcctcctgctcctcggcacaatagcctgc 1108  
DB 1099 TGCAGAACCCAAACCGCAAGGGCCACTGTGGCTCTCTGCTCGGCACTAGCCTGC 1158  
QY 1109 cgtggccgctggtcagtcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1168  
DB 1159 CGTGGCCCGGTGTATCGGATCCACAGCGCGGCTGTCTTCATCTTGAGGCAACCTGTCTCC 1218  
QY 1169 tctggtcttgacagggcccaagaagaagcgctgaccccgccgccccttcctcctgcctg 1228  
DB 1219 TGTGGCTTTGGCCAGGCCCCAGAGAGCCGTGCACCCCGCGCTCCCTCCCTCGCCTG 1278  
QY 1229 ggcacgcccgcgcagggagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1288  
DB 1279 GGCACGCGCGCGCGGGAGCGCGCGCGAGCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1338  
QY 1289 cgcgcctcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1348  
DB 1339 CCGGCTTCAAGGCTGGGCTGTGTGGGGCTGTGTGAAGACATAGGGTCTCCGGAGGCC 1398  
QY 1349 cccgagcacttctgagcccaagcagctgctgcccctaagttgtaaccacaactctaca 1408  
DB 1399 CCCAGCACTTACTGGGCCCAAGGCCCAAGTTGCTGGCCCTTAAGTTGATCCCAACTCTACA 1458  
QY 1409 cagacatccac 1431  
DB 1459 CAGACATCCACACACACACACACAC 1481

RESULT 8  
AK000530  
LOCUS AK000530 759 bp mRNA PRI 22-FEB-2000  
DEFINITION Homo sapiens cDNA FLJ20523 fis, clone KAT10456.  
ACCESSION AK000530  
VERSION AK000530.1 GI:7020687  
KEYWORDS oligo capping; f1s (full insert sequence).  
SOURCE Homo sapiens signet-ring cell carcinoma cell line:KATO III CDNA to mRNA, clone.lib:KAT clone:KAT10456.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (sites)  
AUTHORS Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.  
NEDO human cDNA sequencing project  
TITLE Unpublished (2000)  
JOURNAL 2 (bases 1 to 759)  
REFERENCE Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.  
AUTHORS Direct Submission  
TITLE Submitted (15-FEB-2000) to the DDBJ/EMBL/Genbank databases. Sumio

COMMENT  
Sugano, Institute of Medical Science, University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1, Minato-Ku, Tokyo 108-8639, Japan (E-mail:cdna1@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)  
NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan: cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES  
source  
1..759  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_line="KATO III"  
/cell\_type="signet-ring cell carcinoma"  
/clone="KAT10456"  
/clone.lib="KAT"  
/note="cloning vector pNE18SFL3"  
236..607  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BA91234.1"  
/db\_xref="GI:7020688"  
/translation="MRDFPCSTVTPTPLPASVSPASVPASIPASVPMFRLFPAALGA QESTPAVGMGMGHSPPERLEPMASSSPHASPDPDEKPMWLYLNEDNINNDGRK TGLQGLWSLGLPGRLVFOPC"  
BASE COUNT 172 a 227 c 190 g 170 t  
ORIGIN

Query Match 23.2%; Score 738.4; DB 89; Length 759;  
Best Local Similarity 99.9%; Pred. No. 3,7e-127;  
Matches 739; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2424 atgctctggaacacacacatgagcagatagctgctgagcagacagacacagcagat 2483  
DB 1 ATTTGCCCGGACACACACATGTGCACAGATATGCTGTGCATGACACACAGTGCAGAT 60  
QY 2484 atgctctcgaatacacacagcagcagcagcagcagcagcagcagcagcagcagcagc 2543  
DB 61 ATGCTGTCCGATACACACGACGACGACATGACATATGCTGCTGGGACACACTTCC 120  
QY 2544 ggaacacatgacacacacaggtgacagatagctgctgagcagacacagcagcagcagcagc 2603  
DB 121 GGACACACATGACACACACAGATGACATATGCTGCTGGACACACGAGACTGAGTCT 180  
QY 2604 ttggagaggtgctgcgtggaagctgacagtagctgctgagcagcagcagcagcagcagcagc 2663  
DB 181 TTTGGAGGCTGTGCGGTGAAGCTGCAGTACGTGTGCGGTGAGGCTTATAGTGTAGAG 240  
QY 2664 ggaacttcctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2723  
DB 241 GGACTTTCCTGCTGCACACGCTACCTCCCACTGCGCCGCTCTGTGCCGCTCACT 300  
QY 2724 cccgcctcatalcccgagcctgtgctccctgagcagcagcagcagcagcagcagcagcagc 2783  
DB 301 CCCGCGCTCAATCCCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
QY 2784 ggggtgcccaggaagtcctcctactgctgtggtggtggtggtggtggtggtggtggtggt 2843  
DB 361 GGGTCCCAAGAGTCCCTACTGCTGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 420  
QY 2844 tgaagagctgagccacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2903  
DB 421 TGAAGAGCTGAGAGCCCAAGCTAGTGGCTCAATCCCACTGATTCCTCCCTGACACAGA 480  
QY 2904 gaaggagccttgatattataattagaagaatgaataataataataatgaatgaatgaatgaat 2963  
DB 481 GAAGGGGCTTGGTATTTATATTAAGAAATGAAGATATATATATATATATATATATATATAT 540  
QY 2964 gactggtgtgcaaggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3023

```
|||||
Db 541 GACTGGATTGCAAGGACTGTCTCTCTCGGGCCCGGAGACCCGCTGTTCACGC 600
Qy 3024 atgctatgaccacaccccgctccaggcagacaccccccaaccactgtcgtgtg 3083
Db 601 ATGCTGATGACACACCCCGTCCAGGACACACCCCGACCTGTGCTGTG 660
Qy 3084 cccagatctctgtaatttattatgtagtttagagctggaagcccgatatttatt 3143
Db 661 CCCCAATCTCTGTAATTTATGTAGAGTTTGAAGTGAAGCCCGTATATTTAT 720
Qy 3144 ttgttaacatgaagtga 3163
Db 721 TTGTTAAACATGAAGTGCA 740

RESULT 9
AX080801 781 bp DNA PAT 27-FEB-2001
LOCUS AX080801
DEFINITION Sequence 47 from Patent WO0109327.
ACCESSION AX080801
VERSION AX080801.1 GI:13169773
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 781)
AUTHORS Ashkenazi,A.J., Baker,K.P., Goddard,A., Godowski,P.J., Gurney,A.L.,
Kijavini,I.J., Lafleur,M., Mark,M.R., Marsters,S.A., Pittl,R.M.,
Matanabe,C.K. and Wood,W.I.
METHOD Method of preventing the injury or death of retinal cells and
treating ocular diseases
Patent: WO 0109327-A 47 08-FEB-2001;
JOURNAL
Genentech, Inc. (US)
LOCATION/Qualifiers
FEATURES
1. 781
Source /organism="synthetic construct"
/db_xref="taxon:32630"
/note="Partial length DNA sequence used in the isolation
of DNA52192."
BASE COUNT 147 a 247 c 273 g 105 t 9 others
ORIGIN
Query Match 19.1%; Score 608.6; DB 10; Length 781;
Best Local Similarity 95.6%; Pred. NO. 4.4e-103;
Matches 631; Conservative 0; Mismatches 28; Indels 1; Gaps 1;
```

```
|||||
Db 479 GGGCGGGGTGATGCACAGCCCGGGGTAGCTCGTGGGCTCAAGTGTGGCCAGC 538
Qy 509 ggcacccctcgcccgacatcagctgtagtgaagagacacagcctttagagcccaagg 568
Db 539 GGCACCCCTCGGCCCACTATCACTGATGAAGAGACAGACCCCTTGAGCGCCCAAGG 598
Qy 569 ccgctgagcccaagagaagatgtagacatgagccttgaagacactgagccggaagaga 628
Db 599 CCGCTGAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 658
Qy 629 gcggcaaatacacatcgccgctgctgacacccgagccgagccatcaagccactaagg 688
Db 659 GCGCAAAATACACTGCGCGGTGCAACCGGCGGGGCGCATTAACGCACTTTCAGG 718
Qy 689 tgtagtgcacagcagcagccgttcccaagccgctgctacagagcagcagccggtgaaca 748
Db 719 TGGATGTGATCCAGNGACCCCGTTTAAAGCCCGTTTAAAGGAGAGAGAGAGAGAGAG 778
```

```
RESULT 10
AC079939 165443 bp DNA HTG 19-SEP-2000
LOCUS AC079939
DEFINITION Mus musculus clone RP23-307J9 strain C57BL6/J, WORKING DRAFT
SEQUENCE, 17 unordered pieces.
ACCESSION AC079939
VERSION AC079939.1 GI:10190774
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 165443)
AUTHORS Bouffard,C.G., Dietrich,N.L., Eagle,W.O., Gupta,J., Ho,S.-L.,
Idol,J., Lee-Lin,S.-Q., Legaspi,R., Lim,M., Maduro,Q.L.,
Maduro,V.B., Mastrian,S.D., McCloskey,J.C., McDowell,J.,
Maduro,R., Snyder,B., Stantirpop,S., Summers,T.J., Thomas,J.W.,
Thomas,P.J., Tjongson,E.E., Touchman,J.W., Tran,J.T., Vogt,J.L.,
Walker,M.A., Wetherby,K.D. and Green,E.D.
MISC Mouse Sequencing Initiative
Unpublished
2 (bases 1 to 165443)
Green,E.D.
Direct Submission
Submitted (19-SEP-2000) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.mouse@nhgri.nih.gov
----- Project Information
Center project name: agy
Center clone name: 307J09
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 152784 bases at least Q40
Consensus quality: 156054 bases at least Q30
Consensus quality: 157806 bases at least Q20
Insert size: 16200; agarose-fp
Insert size: 163843; sum-of-ctnigs
Quality coverage: 4.87x in Q20 bases; agarose-fp
Quality coverage: 4.82x in Q20 bases; sum-of-ctnigs
-----
```

\* NOTE: This is a 'working draft' sequence. It currently consists of 17 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as \* runs of N, but the exact sizes of the gaps are unknown. \* This record will be updated with the finished sequence

```
* as soon as it is available and the accession number will
* be preserved.
1
2741: contig of 2741 bp in length
2742 2841: gap of unknown length
2842 5620: contig of 2779 bp in length
5621 5720: gap of unknown length
5721 7880: contig of 2160 bp in length
7881 7980: gap of unknown length
7981 11912: contig of 3932 bp in length
11913 12012: gap of unknown length
12013 17030: contig of 5018 bp in length
17031 17130: gap of unknown length
17131 19758: contig of 2628 bp in length
19759 19858: gap of unknown length
19859 23684: contig of 3826 bp in length
23685 23784: gap of unknown length
23785 27967: contig of 4183 bp in length
27968 34535: contig of 6468 bp in length
34536 34635: gap of unknown length
34636 42708: contig of 8073 bp in length
42709 42808: gap of unknown length
42809 49787: contig of 6979 bp in length
49788 49887: gap of unknown length
49888 61083: contig of 11196 bp in length
61084 61184: gap of unknown length
61184 79632: contig of 18449 bp in length
79633 79732: gap of unknown length
79733 96572: contig of 16840 bp in length
96573 96672: gap of unknown length
96673 116170: contig of 19498 bp in length
116171 138499: contig of 22229 bp in length
138500 138599: gap of unknown length
138600 165443: contig of 26844 bp in length.
Location/Qualifiers
1. 165443
/organism="Mus musculus"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/clone_lib="RP23-307J9"
/clone="RP23-307J9"
1. 2741
misc_feature
1. 2741
/note="assembly_fragment"
2842..3620
misc_feature
/note="assembly_fragment"
5721..7880
misc_feature
/note="assembly_fragment"
7981..11912
misc_feature
/note="assembly_fragment"
12013..17030
misc_feature
/note="assembly_fragment"
17131..19758
misc_feature
/note="assembly_fragment"
19859..23684
misc_feature
/note="assembly_fragment"
23785..27967
misc_feature
/note="assembly_fragment"
28068..34535
misc_feature
/note="assembly_fragment"
34636..42708
misc_feature
/note="assembly_fragment"
42809..49787
misc_feature
/note="assembly_fragment"
49888..61083
misc_feature
/note="assembly_fragment"
61184..79632
misc_feature
/note="assembly_fragment"
79733..96572
misc_feature
/note="assembly_fragment"
96673..116170
misc_feature
/note="assembly_fragment"
clone_end:17

* as soon as it is available and the accession number will
* be preserved.
1
2741: contig of 2741 bp in length
2742 2841: gap of unknown length
2842 5620: contig of 2779 bp in length
5621 5720: gap of unknown length
5721 7880: contig of 2160 bp in length
7881 7980: gap of unknown length
7981 11912: contig of 3932 bp in length
11913 12012: gap of unknown length
12013 17030: contig of 5018 bp in length
17031 17130: gap of unknown length
17131 19758: contig of 2628 bp in length
19759 19858: gap of unknown length
19859 23684: contig of 3826 bp in length
23685 23784: gap of unknown length
23785 27967: contig of 4183 bp in length
27968 34535: contig of 6468 bp in length
34536 34635: gap of unknown length
34636 42708: contig of 8073 bp in length
42709 42808: gap of unknown length
42809 49787: contig of 6979 bp in length
49788 49887: gap of unknown length
49888 61083: contig of 11196 bp in length
61084 61184: gap of unknown length
61184 79632: contig of 18449 bp in length
79633 79732: gap of unknown length
79733 96572: contig of 16840 bp in length
96573 96672: gap of unknown length
96673 116170: contig of 19498 bp in length
116171 138499: contig of 22229 bp in length
138500 138599: gap of unknown length
138600 165443: contig of 26844 bp in length.
Location/Qualifiers
1. 165443
/organism="Mus musculus"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/clone_lib="RP23-307J9"
/clone="RP23-307J9"
1. 2741
misc_feature
1. 2741
/note="assembly_fragment"
2842..3620
misc_feature
/note="assembly_fragment"
5721..7880
misc_feature
/note="assembly_fragment"
7981..11912
misc_feature
/note="assembly_fragment"
12013..17030
misc_feature
/note="assembly_fragment"
17131..19758
misc_feature
/note="assembly_fragment"
19859..23684
misc_feature
/note="assembly_fragment"
23785..27967
misc_feature
/note="assembly_fragment"
28068..34535
misc_feature
/note="assembly_fragment"
34636..42708
misc_feature
/note="assembly_fragment"
42809..49787
misc_feature
/note="assembly_fragment"
49888..61083
misc_feature
/note="assembly_fragment"
61184..79632
misc_feature
/note="assembly_fragment"
79733..96572
misc_feature
/note="assembly_fragment"
96673..116170
misc_feature
/note="assembly_fragment"
clone_end:17

Query Match 13.5%; Score 429.4; DB 76; Length 165443;
Best Local Similarity 78.4%; Pred. No. 1.1e-70;
Matches 564; Conservative 0; Mismatches 76; Indels 79; Gaps 1;

Y 415 gcaagcagcgcttcaaacagccctcaaatagaagcgccggtgatcgagcgccgtg 474
DB 148676 GCAAGCGGCGCTTCAACAGCCCTCCAAATGAGGCGGAGGATGACGCGCGTGTG 148735

Y 475 ggtacgtcgttgccggtcgaagtgcgtgcagcgagcgagccctgcggccgcatcaagtcg 534
DB 148736 GGTACGCTCTTGCGGCTCAAGTGTGTGGCCAGTGGCCAGCCAGCCAGACATCATGTGG 148795

Y 535 atgaagagcagcagcgcttgcagcgccagagcgccgtgtgagccaggaagaagtcg 594
DB 148796 ATGAAGAGATGACAGACGACTTGAGCGCATCTAGAGGCTAGTGAACAGAAAGAACTGG 148855

Y 595 acaactgagctgaagaactcgtgcgagagagacagcgagcaataacacctgcgctgcg 654
DB 148856 ACACTGAGCTTGAAAGAACTGAGGACCTGGAAGACAGTGGCAAGTACAGCTCCGTATCT 148915

Y 655 aaccgcgcggcgccacacagcgccacacataagttgattgattcc----- 700
DB 148916 AACGAGCGCGGTGCCATCAACGCCACCTACAAAGTGAATGATTCGTAAGTGTGGGTC 148975

Y 701 ----- 700
DB 148976 TGTGTGAGACAGGCGCCCGGTGCTTAATACTGTGTCGACATGTTGTTTTCCTTGG 149035

Y 701 -----agcgagcccgcttcaaacgcccgtgtcctcaagcgagcagcccggtgaacagagtc 755
DB 149036 CTAGAGACCGGACCTGTTCAAGCGCTGTGCTCACAGGAGACACACCTTGAAACAAAGGT 149095

Y 756 ggaactgcggggagacacagtccttcacagtcgaagtcgcagcagcagtcgaagcggatg 815
DB 149096 GGACTTGCGTGGGACAAAGCTTCTCCAGTGCAGAGTCCGCAAGTGAAGCTTGAT 149155

Y 816 ccagtgctgaagcgcgtgtgagtcagtcgagcgagcgccagcacacacacacacacacac 875
DB 149156 CCAAGTGCCTAAGCGGCGGTGAGTACGCGCTCCAGAGGAGCCACACACATGATGT 149215

Y 876 gggcgagcagaagttgtgctgcgtccacaggggtgacgctgtgctgcgcgcagcagtc 935
DB 149216 GGGGAGCCAAAGTTGTGTGTTGCCACAGGGGAGTGTGTGTCACAGGCGCTTGAGCTG 149275

Y 936 ctacctataaagtctgtatcaaccgctgcgcgcaggaagatgagcgatgtatcatctc 995
DB 149276 CTACCTCAACAAAGTGTCTCTCGGCCCGCAGAGATGATGCGCATGATCATCTG 149335

Y 996 ccttgagcccaacacagctgagtcacagcttcgagcgcccttcacagctgcgcgcag 1054
DB 149336 CCTTAGGTGCAAAATACCATGAGGCTACAGTTCCGTAAGCGCTTCTCTCATGTATTACAG 149394

RESULT 11
AC019103 168844 bp DNA HTG 17-AUG-2000
LOCUS Homo sapiens chromosome 4 clone RP11-460119, WORKING DRAFT
DEFINITION AC019103
AC019103 GI:9838209
ACCESSION AC019103.5
VERSION HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```







FEATURES	source
* 137599	137698: gap of unknown length
* 137699	143905: contig of 6207 bp in length
* 143906	144005: gap of unknown length
* 144006	148848: contig of 4844 bp in length
* 148850	148949: gap of unknown length
* 148950	153501: contig of 4552 bp in length
* 153502	153601: gap of unknown length
* 153602	155687: contig of 2086 bp in length
* 155688	155787: gap of unknown length
* 155788	157117: contig of 1330 bp in length
* 157118	157217: gap of unknown length
* 157218	158032: contig of 1815 bp in length
* 159033	158132: gap of unknown length
* 159133	160783: contig of 1651 bp in length
* 160784	160883: gap of unknown length
* 160884	162024: contig of 1141 bp in length
* 162025	162124: gap of unknown length
* 162125	163727: contig of 1603 bp in length
* 163728	163827: gap of unknown length
* 163828	164994: contig of 1167 bp in length
* 164995	165094: gap of unknown length
* 165095	166425: contig of 1331 bp in length
* 166426	166525: gap of unknown length
* 166526	167710: contig of 1185 bp in length.
Location/Qualifiers	
1..167710	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/chromosome="3q"	
/clone="RP11-529F4"	
BASE COUNT	42793 a 41021 c 40795 g 41388 t 1713 others
ORIGIN	





AC069450/C	standard; DNA; HTG; 157647 BP.
XX AC069450	
XX AC069450;	
SV AC069450.4	
DT 02-JUN-2000 (Rel. 63, Created)	
DT 28-JUN-2000 (Rel. 64, Last updated, Version 5)	
XX Homo sapiens clone GS3-445D10, WORKING DRAFT SEQUENCE, 15 unordered pieces	
XX HTG: HTGS_DRAFT; HTGS_PHASE1.	
XX Homo sapiens (human)	
OC Eutheria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;	
OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
XX [1]	
RP 1-157647	
RA Metzker M.L., Lewis L.R., Hume J., Edwards C., Harris C., Dedrich D.,	
RA Thomas S., Okunodu G., Carlack C., Garner T., Addison S., Pace A.,	
RA Williams G., Bonin D., Brooks A., Brown J., Bubay C., Bunac C.,	
RA Burkett C., Chacko J., Chen G., Chen Z., Cox C., Davis C., Delgado O.,	
RA Ding Y., Dugan-Rocha S., Fernandez C., Ferraguto D., Forcum-Tansey J.,	
RA Gyll R., Gorrell J.H., Gunaratne P., Haller G., Hernandez J., Hogue M.,	
RA Hosak H., Hou X., Huber J., Jackson L., Jia Y., Kelly J., Kelly S.,	
RA Kovac C., Liu J., Liu W., Louissegh H., Lozano R.J., Martin R., Massey E.,	
RA McLeod M.P., Mel G., Moore S., Morgan M., Morris S., Neal D., Nelson A.,	
RA Nguyen R., Nguyen N., Ogih M., Parish B., Perez L., Reiter D., Say J.,	
RA Shen H., Vasquez L., Watlington S., Williamson A., Wrensford G., Zhou X.,	
RA Bouck J., Hodgson A., Muzny D.M., Rives M., Scherer S., Sodergren E.,	
RA Weinstock G., Worley K., Gibbs R.;	
RT "Direct Submission";	
RL Unpublished.	
XX [2]	
RP 1-157647	
RA Worley K.C.;	
RT Submitted (30-MAY-2000) to the EMBL/GenBank/DBJ databases.	
RL Human Genome Sequencing Center, Department of Molecular and Human Genetics	
RL Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	
XX On Jun 13, 2000 this sequence version replaced gi:8312016.	
CC ----- Genome Center	
CC Center: Baylor College of Medicine	
CC Center code: BCM	
CC Web site: <a href="http://www.hgsc.bcm.tmc.edu/">http://www.hgsc.bcm.tmc.edu/</a>	
CC Contact: hgsc-help@bcm.tmc.edu	
CC ----- Project Information	
CC Center project name: MMAD	
CC Center clone name: GS3-445D10	
CC ----- Summary Statistics	
CC Sequencing vector: M13; L08821	
CC Chemistry: Dye-primer Bodipy; 100% of reads	
CC Assembly program: Phrap; version 0.990329	
CC Consensus quality: 148880 bases at least Q40	
CC Consensus quality: 157291 bases at least Q30	
CC Consensus quality: 160755 bases at least Q20	
CC Estimated insert size: 143469; sum-of-contigs estimation	
CC Quality coverage: 7.3x in Q20 bases; sum-of-contigs estimation	
CC -----	
CC * NOTE: This is a 'working draft' sequence. It currently	
CC * consists of 15 contigs. The true order of the pieces	
CC * is not known and their order in this sequence record is	
CC * arbitrary. Gaps between the contigs are represented as	
CC * runs of N, but the exact sizes of the gaps are unknown.	
CC * This record will be updated with the finished sequence	
CC * as soon as it is available and the accession number will	
CC * be preserved	
CC * 1 96442: contig of 96442 bp in length	
CC * 96443 96542: gap of unknown length	

	*	* 96543	139770:	contig of 43228 bp in length	
CC	**	139771	139870:	gap of unknown length	
CC	**	139871	141364:	contig of 1494 bp in length	
CC	**	141365	141464:	gap of unknown length	
CC	**	141465	143454:	contig of 1390 bp in length	
CC	**	143455	143554:	gap of unknown length	
CC	**	143555	144962:	contig of 1408 bp in length	
CC	**	144963	145062:	gap of unknown length	
CC	**	145063	146595:	contig of 1533 bp in length	
CC	**	146596	146695:	gap of unknown length	
CC	**	146696	147902:	contig of 1207 bp in length	
CC	**	147903	148002:	gap of unknown length	
CC	**	148006	149085:	contig of 1083 bp in length	
CC	**	149086	149185:	gap of unknown length	
CC	**	149186	150263:	contig of 1078 bp in length	
CC	**	150264	150363:	gap of unknown length	
CC	**	150364	151641:	contig of 1278 bp in length	
CC	**	151642	151741:	gap of unknown length	
CC	**	151742	152954:	contig of 1213 bp in length	
CC	**	152955	153054:	gap of unknown length	
CC	**	153055	154093:	contig of 1039 bp in length	
CC	**	154094	154194:	gap of unknown length	
CC	**	154194	155206:	contig of 1013 bp in length	
CC	**	155207	155306:	gap of unknown length	
CC	**	155307	156540:	contig of 1234 bp in length	
CC	**	156541	156640:	gap of unknown length	
CC	**	156641	157647:	contig of 1007 bp in length.	
FH FH FT FT FT XX	Key	Location/Qualifiers			
SO	Sequence	1..157647 /db_xref="taxon:9606" /organism="Homo sapiens" /clone="GS3-445D10"			
		Query Match	Best Local Similarity	Matches	
OY	1779	acacacacagcagacacacactacccttgatgcatgtatgcatacacacatatggcgccagc	4.4%;	Score 139.8; DB 26; Length 157647; Pred. No. 5.8e-17;	
Db	50758	ACATGCCAACAAACACCAACAATCACACAGGGGCTTTAGTCATGCACAAACACACAAACA	Conservative	0; Mismatches 392; Indels 14; Gaps	
OY	1839	tgtctcccgaaggcaacagttagcagaacagcacatgacagatatgccccgttggaca			1898
Db	50698	TACAGAGCCTATTATGTGCACACACACACTTATCACAAAAGGCTATTATGTGCACACTATCACACA			50639
OY	1899	cagataagctgccccaattgcaacgcagcacagcagacagacatgcagacatatcaaagacat			1958
Db	50638	TACACACAAAGGCTATTATGTGCATGACACACACACACAT-----ACCATATAGGGCTTAT			50587
OY	1959	gctgcctgaatacatacacycaaccaacctgycagatgtytcgtccttgayacacaaacac			2018
Db	50586	AGTACATGCACACACACACACTTATACACAGGGGCTTACATGATGCACACACACACACATPAC			50527
OY	2019	acacagatatgctgctcgtggaagcacaacagctgcagatatgcttatcccggaacacacgtygc			2078
Db	50526	ACACAGGGCTTTATAGTGATGCACACACACACTATCACACAGGGGCTATTATGCTATGCACAC			50467
OY	2079	acagatatgctgccttgagacacacagatatgctgccttgacacacacatgatgacagatat			2138
Db	50466	ACACACATATACACACAGGGGCTTACATGTGCATGCACACACACACACACTATACACACAGGGCTATA			50407
OY	2139	gcttggaacacacacacacacgltgtgcacagatatgctgcttttggacacgcaacacatg			2198
Db	50406	GTCATGTGAACCCACACACACATATACACAGAGCTTATTATGTGCACACACACACATATACACA			50347
OY	2199	cagatatgctgccttgagacacacacttcagaacacagctgcacagcgcagatatgctgcc			2258
Db	50346	GGCGATTATGTGCATGCACACACACACTTATACACAGGGGCTTATTATATATGCACACACACA			50287



repeat_region	/rpt_family="MTD"	complement(13321..13598)
repeat_region	/rpt_family="MMERNLTR"	13597..13778
repeat_region	/rpt_family="MTD"	complement(13948..14046)
repeat_region	/rpt_family="MMB1F"	14293..14381
repeat_region	/rpt_family="PB1"	14442..14558
repeat_region	/rpt_family="MMB1"	14598..14673
repeat_region	/rpt_family="B4"	14629..14724
repeat_region	/rpt_family="MMSOR2"	14851..14991
repeat_region	/rpt_family="ORR1B-INT"	14880..15016
repeat_region	/rpt_family="MMB1"	15781..15911
repeat_region	/rpt_family="MMB1F"	complement(15997..16184)
repeat_region	/rpt_family="B3"	16299..16422
repeat_region	/rpt_family="MMB1F"	16429..16643
repeat_region	/rpt_unit=CAAA	16479..16707
repeat_region	/rpt_family="MMB2"	16719..16923
repeat_region	/rpt_family="B3"	complement(17436..17469)
repeat_region	/rpt_family="AT_rich"	complement(17472..17515)
repeat_region	/rpt_unit=CA	complement(17558..17625)
repeat_region	/rpt_family="B4"	17707..17824
repeat_region	/rpt_family="PB1D9"	complement(17892..17976)
repeat_region	/rpt_family="MMSQR2"	complement(18032..18164)
repeat_region	/rpt_unit=GAAGA	complement(18165..18233)
repeat_region	/rpt_family="MMB2"	complement(18229..18274)
repeat_region	/rpt_family="B3"	complement(18327..18499)
repeat_region	/rpt_family="MMB2"	18570..18781
repeat_region	/rpt_family="B4"	18656..18781
repeat_region	/rpt_family="PB1D10"	complement(18791..18827)
repeat_region	/rpt_unit=GAAGA	complement(18828..18972)
repeat_region	/rpt_family="MMB1"	19004..19085
repeat_region	/rpt_family="B4A"	19097..19266
repeat_region	/rpt_family="MMB2"	19303..19402
repeat_region	/rpt_family="B4A"	19402..19551
repeat_region	/rpt_family="MMB1"	19566..19594
repeat_region	/rpt_unit=TTAAA	complement(20527..20633)
repeat_region	/rpt_family="MMB1"	complement(20740..20765)
repeat_region	/rpt_unit=GA	complement(20810..20831)
repeat_region	/rpt_family="AT_rich"	

[illegible]

LOCUS	AF100956	273800 bp	DNA	ROD	03-NOV-1998
DEFINITION	Mus musculus major histocompatibility locus class II region; Fas-binding protein daxx (DAXX) gene, partial cds; B1ng1 (B1ng1), tapasin (tapasin), RalGDS-like factor (RLF), KE2 (KE2), B1NG4 (B1NG4), beta1, 3-galactosyl transferase (beta1.3-galactosyl transferase), ribosomal protein subunit S18 (RPS18), Sacm21 (Sacm21), H2-K1(D) (H2-K1(D)), R1NG1 (R1NG1), KE6a (KE6a), KE4 (KE4), RKBeta (RKBeta), collagen alpha-2 (X1) (COLA11A2), H2-O alpha (H2-Oalpha), R1NG3 (R1NG3), H2-M alpha (H2-M alpha), H2-M beta 2 (H2-M beta2), and H2-M beta1 (H2-M beta1) genes, complete cds; and Imp 2 gene, partial cds.				
ACCESSION	AF100956				
VERSION	AF100956.1 GI:3811374				
KEYWORDS					
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 273800)				
TITLE	Bowen, L., Qiu, S., Madan, A., Loretz, C., James, R., Dots, M., Mix, L., Hall, J., Lasky, S., and Hood, L.				
JOURNAL	Sequence of the mouse major histocompatibility locus class II region				
REFERENCE	Unpublished				
AUTHORS	2 (bases 1 to 273800)				
TITLE	Bowen, L.				
COMMENT	Direct Submission Submitted (26-OCT-1998) Department of Molecular Biotechnology, Box 357730 University of Washington, Seattle, WA 98195, USA This sequence overlaps the entry in Genbank Accession Number AF027865 by 22022 bases. Sequencing methodology: high redundancy shotgun, using M13 templates. Interspersed Repeats were identified with RepeatMasker (available from http://ftp.genome.washington.edu/RM/RepeatMasker.html).				
FEATURES	location/Qualifiers				
source	1..273800				
source	/organism="Mus musculus"				
source	/strain="129svj"				
source	/db_xref="taxon:10090"				
source	/chromosome="17"				
source	1..139758				
source	/organism="Mus musculus"				
source	/strain="129svj"				
source	/db_xref="taxon:10090"				
source	/clone="BAC 445D10"				
source	/clone_lib="Genome Systems"				
source	<1..2000				
source	/gene="DAXX"				
source	join(<146..208,325..544,1825..2000)				
source	/gene="DAXX"				
source	/product="Fas-binding protein Daxx"				
source	join(<146..208,325..544,1825..1884)				
source	/gene="DAXX"				
source	/note="Intron-exon boundaries defined in relation to cDNA sequence in Genbank Accession Number AF006040"				
source	/codon_start=2				
source	/product="Fas-binding protein Daxx"				
source	/protein_id="AAC69891.1"				
source	/db_xref="GI:3811375"				
source	/translation="PPSKRRFETKRGKGLGNSYTKEPYAADSGONTSVQPMPSPLASVAVSDSTVDSPEHETVNSILCSPPSILLQTPQKSLRRCIYTSVATQCDPEEIVLSDSD"				
source	724..858				
source	/rpt_family="MMB1"				
source	complement(1257..1297)				
source	/rpt_family="(CA)n"				
source	1303..1421				
source	/rpt_family="MMB1.F"				
source	complement(1426..1630)				
source	/rpt_family="B3"				
source	complement(1646..1792)				
source	/rpt_family="MMB1.F"				
source	complement(2728..2788)				

[illegible]



```
----- Summary Statistics -----
Sequencing vector: M13; 96%
Sequencing vector: plasmid; 1%
Chemistry: Dye-primer ET; 96% of reads
Chemistry: Dye-terminator Big Dye; 1% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 164931 bases at least Q40
Consensus quality: 171635 bases at least Q30
Consensus quality: 175162 bases at least Q20
Insert size: 196000; agarose-fp
Quality coverage: 5.65 in Q20 bases; agarose-fp
Quality coverage: 5.00 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1792: contig of 1792 bp in length
* 1793 1892: gap of unknown length
* 1893 4419: contig of 2527 bp in length
* 4420 4519: gap of unknown length
* 4520 6906: contig of 2387 bp in length
* 6907 7006: gap of unknown length
* 7007 11504: contig of 4498 bp in length
* 11505 11604: gap of unknown length
* 11605 17234: contig of 5630 bp in length
* 17235 17334: gap of unknown length
* 17335 22640: contig of 5306 bp in length
* 22641 22740: gap of unknown length
* 22741 29954: contig of 7214 bp in length
* 29955 30054: gap of unknown length
* 30055 37086: contig of 7032 bp in length
* 37087 43364: gap of unknown length
* 43365 43464: gap of unknown length
* 43465 53529: contig of 10065 bp in length
* 53530 53629: gap of unknown length
* 53630 61924: contig of 8295 bp in length
* 61925 62024: gap of unknown length
* 62025 72580: contig of 10556 bp in length
* 72581 72680: gap of unknown length
* 72681 86002: contig of 13322 bp in length
* 86003 86102: gap of unknown length
* 86103 96656: contig of 10554 bp in length
* 96657 96756: gap of unknown length
* 96757 111536: contig of 14780 bp in length
* 111537 111636: gap of unknown length
* 111637 127101: contig of 15465 bp in length
* 127102 127201: gap of unknown length
* 127202 141174: contig of 13973 bp in length
* 141175 141274: gap of unknown length
* 141275 156588: contig of 15314 bp in length
* 156589 156688: gap of unknown length
* 156689 157569: contig of 881 bp in length
* 157570 157669: gap of unknown length
* 157670 186583: contig of 28914 bp in length.

FEATURES
source
1..186583
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="7"
  /clone="RP11-369D24"
  1..1792
    /note="assembly_name:Contig27"
    1893..4419
      /note="assembly_name:Contig28"
      4520..6906
        /note="assembly_name:Contig29"
```

```
misc_feature 7007..11504
  /note="assembly_name:Contig30"
misc_feature 11605..17234
  /note="assembly_name:Contig31"
misc_feature 17335..22640
  /note="assembly_name:Contig32"
misc_feature 22741..29954
  /note="assembly_name:Contig33"
misc_feature 30055..37086
  /note="assembly_name:Contig34"
misc_feature 37187..43364
  /note="assembly_name:Contig35"
misc_feature 43465..53529
  /note="assembly_name:Contig36"
misc_feature 53630..61924
  /note="assembly_name:Contig37"
misc_feature 62025..72580
  /note="assembly_name:Contig38"
misc_feature 72681..86002
  /note="assembly_name:Contig39"
misc_feature 86103..96656
  /note="assembly_name:Contig40"
misc_feature 96757..111536
  /note="assembly_name:Contig41"
misc_feature 111637..127101
  /note="assembly_name:Contig42"
misc_feature 127202..141174
  /note="assembly_name:Contig43"
misc_feature 141275..156588
  /note="assembly_name:Contig44"
  /vector_end="sp6"
  /vector_side="left"
misc_feature 156689..157569
  /note="assembly_name:Contig45"
misc_feature 157670..186583
  /note="assembly_name:Contig45"
BASE COUNT 41658 a 49671 c 52352 g 40786 t 1916 others
ORIGIN
Query Match 4.3%; Score 136.6; DB 74; Length 186583;
Best Local Similarity 53.9%; Pred. No. 2.2e-16;
Matches 444; Conservative 0; Mismatches 354; Indels 25; Gaps 7;
QY 1778 cacacacacagacacacacactactcgtgacgtatgtatgtacacacacatgcgcgcacac 1837
  |||||
Db 125988 CACACACACAGACACGACACATCAGATGCACCCACACAGACGACACATCAGATGCA 125929
  |||||
QY 1838 gtctccctgaagcgacagctgcgcacacacacacacacacacacacacacacacacacac 1897
  |||||
Db 125928 CCCACACAGACACGACACATGACCCACACAGACACGACACATCAGATGCA 125869
  |||||
QY 1898 acagataagctgcgaatgcacacacacacacacacacacacacacacacacacacacacacac 1956
  |||||
Db 125668 AGACACGACACATCAGATGACACACACACACACACACACACACACACACACACACACAC 125809
  |||||
QY 1957 atgctgcctgaacatacaacacacacacacacacacacacacacacacacacacacacacac 2016
  |||||
Db 125808 ACACACACATCAGATGCACCCACATCAGACACGACACATCAGATGTC---ACACACACAG 125753
  |||||
QY 2017 aacacacgatatgtctgtcggacacacacacacacacacacacacacacacacacacacacacac 2076
  |||||
Db 125752 ACACACACATCAGATGCACACACACACACACACACACACACATCAGATGCA 125693
  |||||
QY 2077 gcaacagatatgctgcctgcggacacacacacacacacacacacacacacacacacacacacacacac 2136
  |||||
Db 125692 GCACACATCAGATGCACCCACACACACACACACACACACACACACACACACACACACAC 125636
  |||||
QY 2137 ttgcctgcacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 2196
  |||||
Db 125635 CACACATCAGATGCACCCACACACACACACACACACACACATCAGATGCA 125576
  |||||
QY 2197 ttgcagatatgctgcctgcggacacacacacacacacacacacacacacacacacacacacacac 2256
  |||||
```

Db	125575	CACATGCACCCACACAGACGCACACATGCAGATGAGTGCACCCACACAGACGACACATCA	125516
Oy	2257	ccctggacacacagcgatattgtctctctgtcaacacacacagcagaaatgtcttcggaca	2316
Db	125515	GATGCACACACACAGACACCCCATGCAGATGCACACACAGAACCCAC-ACATGCACC	125457
Oy	2317	cacacacgcatgacagaatactgtcttcggacacacacacgacacagatatgtctgctg	2376
Db	125456	CACACAGACGCGCACATCTAGATGCACCCACACAGACACGCCACACATCTCAGATGCACACA	125397
Oy	2377	gacacacacagataaattgtctgtcctaactcaacacgltgcagatatgtcttcggac	2436
Db	125396	CACAGACCCCATCTCA-GATGCACACACAGACACACACATCTCAGATGCACCCATAC	125388
Oy	2437	aacacgtgcagatatgtctgtgcatactgcacacacgltgcagatatgtcttcggat	2496
Db	125337	AGACTGTGCACAT-----CAGATGCACACACAGACACACATCTGAT	125292
Oy	2497	aacacgcacgcacacatatgtcgtctgtgcgcacacacttcggaca-cacatgc	2555
Db	125291	GCACACACACAGACACACACATCAGATGCACCCACACAGACACGCACACATCAGATGC	125232
Oy	2556	acacacaggtgcagatatgtctgtcctgcgcacacacgacagcagatgac	2588
Db	125231	ACACACACGACACACACATCAGATGCACACACACACACAC	125189

RESULT	20
AC027728/c	LOCUS
AC027728	174473 bp DNA HTG 02-MAY-2000
DEFINITION	Homo sapiens chromosome 1 clone RP11-18H19 map 1, WORKING DRAFT SEQUENCE, 14 unordered pieces.
AC027728	AC027728
AC027728.2	GI:7677838
HTG:	HTGS_PHASE1; HTGS_DRAFT.
SOURCE	human.
ORGANISM	Homo sapiens .
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 174473)
TITLE	Birren,B., Linton,L., Nusbaum,C. and Lander,E.
JOURNAL	Homo sapiens chromosome 1, clone RP11-18H19
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 174473)
	Birren,B., Linton,L., Nusbaum,C., Lander,E., Agham,H., Allen,N., Anderson,S., Baldwin,J., Baran,N., Baatlen,Y., Beda,F., Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G., Campiolano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D., Galeana,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Larocque,K., Lamazars,R., Landers,T., Lechoczky,J., Levine,R., Lien,C., Liu,G., Locke,K., MacDonald,P., Margulis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McNeeters,R., Meldrum,J., Menous,L., Mihova,T., Miranda,C., Menga,V., Morrow,J., Murphy,T., Naylor,J., Norma,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M., Olivier,J., Peterson,K., Pierre,N., Pisanic,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talmas,J., Teefegre,S., Theodore,J., Tirelli,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.
TITLE	Direct Submission
JOURNAL	Submitted (01-Apr-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On May 2, 2000 this sequence version replaced gi:7382541. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) <a href="http://ftp.genome.washington.edu/RV/RepeatMasker.html">http://ftp.genome.washington.edu/RV/RepeatMasker.html</a>

```

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: I7571
Center clone name: 18_H_19
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 166277 bases at least Q40
Consensus quality: 170318 bases at least Q40
Consensus quality: 171952 bases at least Q20
Insert size: 175000; agarose-1p
Insert size: 173173; sum-of-contigs
Quality coverage: 4.8 in Q20 bases; agarose-1p
Quality coverage: 4.9 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 493: contig of 493 bp in length
* 494 593: gap of 100 bp
* 594 3150: contig of 2537 bp in length
* 3131 3230: gap of 100 bp
* 3231 6669: contig of 3439 bp in length
* 6670 6769: gap of 100 bp
* 6770 13955: contig of 7186 bp in length
* 13956 14055: gap of 100 bp
* 14056 22245: contig of 8190 bp in length
* 22246 22345: gap of 100 bp
* 22346 30691: contig of 8346 bp in length
* 30692 30791: gap of 100 bp
* 30792 41365: contig of 10574 bp in length
* 41366 41465: gap of 100 bp
* 41466 52077: contig of 10612 bp in length
* 52078 52177: gap of 100 bp
* 52178 66297: contig of 14120 bp in length
* 66298 66397: gap of 100 bp
* 66398 83577: contig of 17180 bp in length
* 83578 83677: gap of 100 bp
* 83678 100579: contig of 16502 bp in length
* 100580 100679: gap of 100 bp
* 100680 121055: contig of 20376 bp in length
* 121056 121155: gap of 100 bp
* 121156 147315: contig of 26160 bp in length
* 147316 147415: gap of 100 bp
* 147416 174473: contig of 27058 bp in length.
*
Location/Qualifiers
1. 174473
/db_xref="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1"
/clone="RP11-18H19"
/clone.lib="RPC1-11 Human Male BAC"
1. 493
/note="assembly-fragment"
clone_end:SP6
vector_side:left"
594. .3130
/note="assembly-fragment"
3231. 6669
/note="assembly-fragment"
6770. .13955
/note="assembly-fragment"

```









```

/map="14g32"
/clone="BAC2B24"
repeat_region complement(49.335)
/rpt_family="AluDb"
repeat_region 451.747
/rpt_family="AluSg"
repeat_region complement(787.963)
/rpt_family="MLT1A2"
repeat_region complement(2213.2414)
/rpt_family="MLT1A1"
repeat_region complement(3658.4664)
/rpt_family="LTR5"
repeat_region complement(5010.5102)
/rpt_family="MER58B"
mRNA
join(5238.5411,21222.21253,25192.25264,36341.36490,
45477.45570,47630.47756,48100.48139,48229.48321,
52476.52742,54260.54321,55158.55239,55337.55392,
56590.56846,62226.62277,63097.63268,63802.63844)
/note="Intron Exon boundaries were defined in relation to
human transcription factor TRF1B 90 kDa subunit sequence
present in 028838"
repeat_region complement(6962.7265)
/rpt_family="AluSx"
repeat_region 7298.7596
/rpt_family="AluY"
8264.8377
/rpt_family="L1MB8"
repeat_region 10084.10140
/rpt_family="L1ME3A"
repeat_region 10159.10463
/rpt_family="AluDb"
repeat_region complement(10710.11009)
/rpt_family="AluDb"
repeat_region complement(11352.11490)
/rpt_family="AluDb"
repeat_region complement(11491.11608)
/rpt_family="AluDb"
repeat_region complement(11684.11852)
/rpt_family="AluSx"
repeat_region complement(11853.12158)
/rpt_family="AluY"
repeat_region complement(12159.12303)
/rpt_family="AluSx"
12429.12711
/rpt_family="AluDb"
repeat_region complement(12915.13075)
/rpt_family="AluDb"
repeat_region complement(13083.13386)
/rpt_family="AluY"
repeat_region complement(13449.13572)
/rpt_family="AluDb"
repeat_region complement(13673.13753)
/rpt_family="(CAAA)n"
repeat_region complement(13825.14125)
/rpt_family="AluY"
repeat_region complement(14514.14811)
/rpt_family="AluSg"
repeat_region complement(14814.14891)
/rpt_family="FLAM_A"
17099.17395
/rpt_family="AluSx"
repeat_region complement(17710.17834)
/rpt_family="L1ME2"
repeat_region complement(17849.18139)
/rpt_family="AluSx"
repeat_region complement(18171.18310)
/rpt_family="L1MD1"
repeat_region complement(18311.18607)
/rpt_family="AluDb"
repeat_region complement(18610.18792)
/rpt_family="L1MD2"
repeat_region complement(18805.18951)
/rpt_family="L1MB8"

repeat_region complement(18961.19254)
/rpt_family="AluSx"
repeat_region complement(19265.19564)
/rpt_family="AluY"
repeat_region complement(19566.19992)
/rpt_family="L1MC2"
repeat_region complement(21627.21926)
/rpt_family="AluDb"
22910.23210
/rpt_family="AluSg"
repeat_region 23215.23515
/rpt_family="AluY"
23778.23881
/rpt_family="(CATG)n"
repeat_region complement(25425.25575)
/rpt_family="AluDb"
repeat_region complement(28632.28681)
/rpt_family="GC_rich"
28752.28878
/rpt_family="(CGG)n"
repeat_region complement(28904.28988)
/rpt_family="GC_rich"
repeat_region complement(29092.29211)
/rpt_family="(CGGG)n"
29239.29289
/rpt_family="GC_rich"
repeat_region complement(29639.29754)
/rpt_family="MIR"
30225.30346
/rpt_family="L1"
30731.30855
/rpt_family="L1ME2"
31011.31276
/rpt_family="AluSx"
repeat_region complement(31277.31368)
/rpt_family="AluDb"
repeat_region complement(31391.31527)
/rpt_family="L1MB6"
repeat_region complement(31532.31665)
/rpt_family="AluSx"
repeat_region complement(31675.32234)
/rpt_family="L1"
repeat_region complement(32270.32380)
/rpt_family="(TAAA)n"
repeat_region complement(32381.32684)
/rpt_family="AluSx"
repeat_region complement(33305.33604)
/rpt_family="AluDb"
repeat_region complement(33610.33736)
/rpt_family="FLAM_C"
repeat_region complement(33747.34050)
/rpt_family="AluSx"
repeat_region complement(34136.34268)
/rpt_family="AluDb"
repeat_region complement(34284.34600)
/rpt_family="AluSx"
repeat_region complement(34626.34925)
/rpt_family="AluSx"
repeat_region complement(35000.35300)
/rpt_family="AluSp"
repeat_region complement(35455.35600)
/rpt_family="L1MA2"
repeat_region complement(45831.46115)
/rpt_family="AluSg"
repeat_region complement(46952.47252)
/rpt_family="AluSg"
repeat_region complement(47255.47552)
/rpt_family="AluDb"
repeat_region complement(48348.48445)
/rpt_family="MER4D"
48447.48742
/rpt_family="AluSg"
repeat_region complement(48743.48879)

```

Db	58420	CACACACTGCATTACACAGGCGCACACACACATGACACACACACAGCTGCAGGCATTACAGACA	5836
QY	2413	caacgtgcagataattgacctgagacacacacatgtgcacagataatgtctgtctgacatgcaca	2472

\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the runs between them are based on estimates that have

```

* Provided by the submittor
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 9835 9934: contig of 9834 bp in length
*      1
* 9935 466472: contig of 36538 bp in length
* 46473 46572: gap of 100 bp
* 46573 189418: contig of 142846 bp in length.

```

```

AGNSIGLSYOSAMTLVLPEDPTWTAAPPEARITDILVAGSLALAVLLIAGLYG
OALHGRHPRPATVOKLSRFLAROFSLSSGSSSSSLVGRVRLSSSGALLAGLY
SLDLPDLMPEPRDLVGLGEGCGOVVRAEAGMDPAQDASTAVKMLKDN
ASDDIADIVSEMEVMTLGRHKNIINILGVCTOEGPLVYVECAAGNLFREFLAR
PPGPPLSDGRRSSRSPULSFPVIVSCAVOVARGMOLYRECRICHDYLAARVLEEN
VVKIADFGIARVHHIDYTKTSNRLPYKMAEALFDPRYTHQSDVMSGILIMEI
FTLGSPYGPILVEELFSLRGRHMDRPPHCPPELIGLMECHMAAPSDRPTKOLY
EALDKVLAVSEBYDLRLTFGPYSPSGDASTSCSSDSVSHDPLPLGSSSPFGS
GVOY"

misc_difference
/1060
/gene="FGFR4"
/Note="Asp in translation in B5/589 Clone 44 vs. Val in
Ref.2.; a in B5/589 clone44 vs. t in [2].; G00-127-929"
/citation=[1]
/replace="t"
/replace="c"
misc_difference
2931..2933
/Note="FGFR4"
/Note="GCG in B5/589 clone 44 vs. CCG in [2].;
G00-127-929"
/citation=[1]
/replace="cgc"
/replace="cgc"

BASE COUNT      540 a      971 c      944 g      576 t

ORIGIN
Query Match      3.9%; Score 123; DB 97; Length 3031;
Best local Similarity 49.7%; Pred. No. 3.1e-13;
Matches 423; Conservative 0; Mismatches 395; Indels 33; Gaps 3;

372 cagctctcctgggggtcaagaagaccgccagccagctggcgacagccgctcac 431
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
560 CAACTCCCATAGGAGACCCCTCGAATAGGACATTTACCCACAGACACCCCTAC 619
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
432 acagccctcaagaatgagcgcggtgatacagcgccgctggtgtagctcgcgct 491
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
620 ACACCCCGGAGCATGAGAGAAAGTGCATGACCTCGCGGGGAAACCTCAAGTT 679
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
492 caagtgtgtggcgagcgagccctcgccagcatcagctgtagtaagaagacagc 551
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
680 CCGCTGTCCAGCTCAGACCAACCCACGCCACCATCCGCTGCTTAAGATGACAGC 739
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

552 ctt-----gaagcgcccaagagcgcttgagcccaagaagaagatggaactgag 602
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
740 CTTTCATGGGGAGAAACCGCATTTGAGGACATTCGGCTCGCATCAGACATCGT 799
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

603 cctgaagaacctgcgcgcgagagagcaagcaataactctgcgcgctgtcgaaacgcgc 662
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
800 GATGAGAGCCTGTGCTCGCTCGACCGCGGCACATACCTGCTGTAAGAAAGCTGT 859
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

663 ggggccaataagccacactcaaggttgatgatccagcgagaccgctccaaagccgt 722
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
860 GGGGAGCATTCGGCTATTAACCTGCTAGATGCTGTGAGGGGCTCCCGGACCGGCCAT 919
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

723 gctcaagagcaagaccgcgtgaacacgagctggaactcggggggaacacgctctcca 782
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
920 COTGAGGCGGGGCTCCCGGCAACACACCGCTGTGTTGGGCGACGACGTTGAGTGTCT 979
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

783 gtgcaaggtgtgcgcgcgtgaaagcggtgatactcaagtgtgtgaagcgctgtgagta 842
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
980 GGGCAAGGTGTAACACCGGTGCGCCAGCCACATCCAGTGCTGAAGC-----ACAT 1030
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

843 ggcgagggcgccgacacatcatcatgagctgagcgccagagaagttgtgtgtgctc 902
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1031 CGTTCATCAACGCGACACTTCGAGCGGAGGTTTCCCTCATGTGCAAGGCTTAAGAC 1090
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

903 cagcggtgacgtgtgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 962
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1091 TGCAGACATCATATAGCTAGAGGTGAGGTCTCTGCTGCGGAACGTTGACGCC----- 1145
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

963 tgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1022
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1146 -----GAGGAGCGAGGCGAGTACACCTGCTCGACGACATTCATCGGCTCTC 1195
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 1023 ctccgcagcgccctctcaccgtgtgcagagaccacacgcgcagggccacctgtgac 1082
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1196 CTACAGCTGTGCTGCTCACGGTGTGCTCCAGAGAGAGACCCACATGACCCGACAGC 1255
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1083 ctccgtcctgcgcgcacacacacacacacacacacacacacacacacacacacacac 1142
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1256 GCCGAGGCGCAGTATACGACATCATCTGTACGCTGCGGCTCCCTGCTGCTGTGT 1315
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1143 tgtcttcacctgtgagacacctgcctcctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1202
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1316 GCTCTGCTGCTGCGCGGCGGTATGACAGGCGAGCGGCTCCACGCGCGGACCCGCCCC 1375
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1203 ccccgcgctg 1213
||| ||| |||
Db 1376 GCCCGCCACTG 1386

RESULT 26
HSFGFR4
LOCUS
DEFINITION
ACCESSION
X57205
VERSION
X57205.1 GI:31371
KEYWORDS
FGF receptor gene; FGFR-4 gene; fibroblast growth factor receptor.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2915)
REFERENCE
1 Partanen,J.M.
Direct Submission
Submitted (14-JAN-1991) J.M. Partanen, UNIVERSITY OF HELSINKI,
DEPARTMENTS OF VIROLOGY & PATHOLOGY, CANCER BIOLOGY LABORATORY,
HAARTMANINKATU 3, 00290 HELSINKI 29, FINLAND
2 (bases 1 to 2915)
Partanen,J., Makela,T.P., Berola,E., Korhonen,J., Hirvonen,H.,
Classon-Weish,L. and Alitalo,K.
FGFR-4, a novel acidic fibroblast growth factor receptor with a
distinct expression pattern
EMBO J. 10 (6), 1347-1354 (1991)
91224085
JOURNAL
MEDLINE
COMMENT
See M59373 for overlapping sequences.
FEATURES
source
location/Qualifiers
1..2915
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..2915
/gene="FGFR-4"
1..2915
/gene="FGFR-4"
56..2464
/gene="FGFR-4"
/codon_start=1
/product="fibroblast growth factor receptor"
/protein_id="CAA40490.1"
/db_xref="GI:31372"
/db_xref="SWISS-PROT:P22455"
/translation="MRLIALGLVLPVGPVYLSLESEVLEPCLAPSLDEOE
LTVAGVRLCCGAEGRGHWYKSGSLAPAGRVGRGRGLTSLFPEAGRTCL
ARGSMIVLONTLITGDSLTSSNDEDESHSDPNRSYDQARVYTHHPOMERKTL
AVPAGNTVAFRCPPAGNPPTIRMLKDCQAFGEENRIGIRLRHQMSLVESVPSD
RGTYCIENANGSIRYNYLIDVERSHRPTLQGLPANTYALRNVAEDAGETCL
AOPHIOMLKHIVYNSSRGANGFPVVOYLKTADINSSEVYLYLRNVAEDAGETCL
AGNSIGLSYOSAMTLVLPEDPTWTAAPPEARITDILVAGSLALAVLLIAGLYG
OALHGRHPRPATVOKLSRFLAROFSLSSGSSSSSLVGRVRLSSSGALLAGLY
SLDLPDLMPEPRDLVGLGEGCGOVVRAEAGMDPAQDASTAVKMLKDN
ASDDIADIVSEMEVMTLGRHKNIINILGVCTOEGPLVYVECAAGNLFREFLAR
PPGPPLSDGRRSSRSPULSFPVIVSCAVOVARGMOLYRECRICHDYLAARVLEEN
VVKIADFGIARVHHIDYTKTSNRLPYKMAEALFDPRYTHQSDVMSGILIMEI
FTLGSPYGPILVEELFSLRGRHMDRPPHCPPELIGLMECHMAAPSDRPTKOLY
EALDKVLAVSEBYDLRLTFGPYSPSGDASTSCSSDSVSHDPLPLGSSSPFGS
GVOY"
BASE COUNT      516 a      941 c      892 g      566 t

```













[illegible]

REFERENCE	ARTIFICIAL SEQUENCE.
ATTORNS	1 (bases 1 to 2427)
TITLE	Cappelletti, D., Chopin, D., Radanyi, F., Ricol, D., and Thery, J.P.
JOURNAL	Means for detecting and treating pathologies linked to fgfr3
	Patent: WO 0068424-A 4 16-NOV-2000;
	Institut Curie (FR) ; CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
	(CNRS) (FR)
FEATURES	Location/Qualifiers
SOURCE	1..2427
	/organism="synthetic construct"
	/db_xref="taxon:32630"
	/note="Mutant G372C FGFR3-T11D."
BASE COUNT	428 a 784 c 820 g 395 t
ORIGIN	
Query Match	3.7%; Score 118.6; DB 9; Length 2427;
Best Local Similarity	56.1%; Pred. No. 2.2e-12;
Matches 270; Conservative	0; Mismatches 199; Indels 12; Gaps 2;
OY	421 ccgcgtccacacagccctcccaagatgtgagcgccggtgcatgcacgcccgttggtagc 480
Db	451 CCTTCTGTGACACGGCCCGAGCGATGGAACAAGAAGCTGCTGGCCGTGGCCGCCACAC 510
OY	481 tccgtgcggtcaagtgtcgttgcacgcgagccctcgcccgccacatcacgtgtgaaga 540
Db	511 ACCGTCGCGCTTCGCTGCCACGCCCTGGCAACCCACCTCCCTCATCTCCTGGCTGAAG 570
OY	541 gacgcacagccct-----gacgcgcccagaagcgcgctgagcccaagaagaag 591
Db	571 AACGGCAGGAGATTCGGCGGCGACGACCGCATTTGAGAGCATCAACCTGGCATCAGCAG 630
OY	592 tggacactgagactgtgaagaacttcgcccgcgagcagcgagcaataaccctgcgcgtg 651
Db	631 TGGAGCTGTGTATGGAAGAGCTGTGCTCCCTCGGACCGGGCACTACACCTGCGTCTG 690
OY	652 tcgaacgcgcgcgagccatcaacgcacactcaagaagtgtgattgtatccagcgaccgt 711
Db	691 GAGAACAAATTTGGCAGAGCTCCGCGAGAGAGTACAGGCTGGAGCTGGAGGCGCTCCCG 750
OY	712 tccaagcccggtgtctcaacgagcaaccccgctggaacaagaagctggaacttcggggagcc 771
Db	751 CACCGGCCCATCTTCGACGCGGGGCTGCGCGCCACACACACGCGGCTGTGGGACGAC 810
OY	772 agctccctccagtcagtcaggtgtgcacgcgagctgtgaagtcggtgctgtgaagcgc 831
Db	811 GGGGAGTTCCATCGCAAGGTGTACAGTACGACCAACCCCATCATCGATGCGTCAAGCAC 870
OY	832 gtgga---gtacggcgccgaagggccgcacaaactccacatcatgatgttggcgccgaag 888
Db	871 GTGGAGGTGAACGGCAGCAAGTGGGCGCGAGGCGACACCTACCTTACCTGCTCMAAG 930
OY	889 t 889
Db	931 T 931
RESULT 34	
LOCUS	AX047600 2427 bp DNA PAT 15-DEC-2000
DEFINITION	Sequence 5 from Patent WO0068424.
ACCESSION	AX047600
VERSION	AX047600.1 GI:11876668
KEYWORDS	
SOURCE	synthetic construct.
ORGANISM	synthetic construct
	artificial sequence.
REFERENCE	1 (bases 1 to 2427)
AUTHORS	Cappelletti, D., Chopin, D., Radanyi, F., Ricol, D., and Thery, J.P.
TITLE	Means for detecting and treating pathologies linked to fgfr3
JOURNAL	Patent: WO 0068424-A 5 16-NOV-2000;
	Institut Curie (FR) ; CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
	(CNRS) (FR)



```

Db 451 CCTTACTGACACGGCCCGAGCGGATGGACAAGAGTGTCTGGCTGCGCCGCGCCAC 510
QY 481 tccgtgagctcaagtgtggtgagcgagccctcgccagcatcaagtgtgagaa 540
Db 511 ACCGTCCGCTTCCGCTGCGACCGCCTGTGACACCCCACTCCCTCATCTCTGCTGAAG 570
QY 541 gacgacagcgctt-----gacgagccagagcgccgtgtgagccaggaagaag 591
Db 571 AACGGAGGAGGATTCGCGCGGAGACCGCATTTGGAGCATCAACCTCGGCATCAGCAG 630
QY 592 tggacaacttgagccttgaagaactctgcccggagagcagcggaataaactcgccgctg 651
Db 631 TGGAGCTGTGATGTAAGAAAGCTGTGTGCTTGGACCGGGCACTTACACTGCTGCTG 690
QY 652 tcgaacccgagcgccatcaacgccaactcaaggtgagatgtatccagcgagccgt 711
Db 691 GAGAACAAATTTGGCAGCATTCGGGAGAGCATGACAGCTGAGCTGAGAGCGCTCCCG 750
QY 712 tccaagccgtgtcacaagcgacgcccgtgaaacagagcggtgtgactctgggggagc 771
Db 751 CACCGGCCCATCTCTGAGCGGCGGCTGCGGCCAACCAACAGCGGCTGCTGGGCAAGCAG 810
QY 772 agctcttcagtgcaaggtgtgcaagcgacgagtgaaagcggtgtgactctgggggagc 831
Db 811 GTGAGGTTCACCTGCAAGGTGTACAGTGAACGACACAGCCCACTCAGTGTGCTCAAG 870
QY 832 gtgga---gtacggcgccgagagcgccacacatccacatcgatgtggcgagcaag 888
Db 871 GTGAGGTGAACGCGCAGCAAGGTGGCCGCGGACGACACCTTACGTTACGTTGCTCAAG 930
QY 889 t 889
Db 931 T 931

RESULT 37
AX047603 2427 bp DNA PAT 15-DEC-2000
LOCUS AX047603 Sequence 8 from Patent WO0068424.
DEFINITION AX047603
ACCESSION AX047603
VERSION AX047603.1 GI:11876671
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 2427)
AUTHORS Cappellen,D., Chopin,D., Radvanyi,F., Ricol,D. and Thierly,J.P.
TITLE Means for detecting and treating pathologies linked to Igfr3
JOURNAL Patent: WO 0068424-A 8 16-NOV-2000;
Institut Curie (FR) ; CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
(CNRS) (FR)
FEATURES
source Location/Qualifiers
1..2427
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Mutant K652M FGFR3-IIIB:"
BASE COUNT 427 a 784 c 821 g 395 t
ORIGIN
Query Match 3.7%: Score 118.6; DB 9; Length 2427;
Best Local Similarity 56.1%: Pred. No. 2.2e-12;
Matches 270; Conservative 0; Mismatches 199; Indels 12; Gaps 2;
QY 421 ccgagcttcacacagcctccaagatgagcgccggtgatcgacgagcccggtgagtc 480
Db 451 CCTTACTGACACGGCCCGAGCGGATGGACAAGAGTGTCTGGCTGCGCCGCGCCAC 510
QY 481 tccgtgagctcaagtgtggtgagcgagccctcgccagcatcaagtgtgagaa 540
Db 511 ACCGTCCGCTTCCGCTGCGACCGCCTGTGACACCCCACTCCCTCATCTCTGCTGAAG 570
QY 541 gacgacagcgctt-----gacgagccagagcgccgtgtgagccaggaagaag 591
Db 571 AACGGAGGAGGATTCGCGCGGAGACCGCATTTGGAGCATCAACCTCGGCATCAGCAG 630
QY 592 tggacaacttgagccttgaagaactctgcccggagagcagcggaataaactcgccgctg 651
Db 631 TGGAGCTGTGATGTAAGAAAGCTGTGTGCTTGGACCGGGCACTTACACTGCTGCTG 690
QY 652 tcgaacccgagcgccatcaacgccaactcaaggtgagatgtatccagcgagccgt 711
Db 691 GAGAACAAATTTGGCAGCATTCGGGAGAGCATGACAGCTGAGCTGAGAGCGCTCCCG 750
QY 712 tccaagccgtgtcacaagcgacgcccgtgaaacagagcggtgtgactctgggggagc 771
Db 751 CACCGGCCCATCTCTGAGCGGCGGCTGCGGCCAACCAACAGCGGCTGCTGGGCAAGCAG 810
QY 772 agctcttcagtgcaaggtgtgcaagcgacgagtgaaagcggtgtgactctgggggagc 831
Db 811 GTGAGGTTCACCTGCAAGGTGTACAGTGAACGACACAGCCCACTCAGTGTGCTCAAG 870
QY 832 gtgga---gtacggcgccgagagcgccacacatccacatcgatgtggcgagcaag 888
Db 871 GTGAGGTGAACGCGCAGCAAGGTGGCCGCGGACGACACCTTACGTTACGTTGCTCAAG 930
QY 889 t 889
Db 931 T 931

RESULT 38
AX047604 2427 bp DNA PAT 15-DEC-2000
LOCUS AX047604 Sequence 9 from Patent WO0068424.
DEFINITION AX047604
ACCESSION AX047604
VERSION AX047604.1 GI:11876672
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 2427)
AUTHORS Cappellen,D., Chopin,D., Radvanyi,F., Ricol,D. and Thierly,J.P.
TITLE Means for detecting and treating pathologies linked to Igfr3
JOURNAL Patent: WO 0068424-A 9 16-NOV-2000;
Institut Curie (FR) ; CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
(CNRS) (FR)
FEATURES
source Location/Qualifiers
1..2427
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Mutant X809C FGFR3-IIIB:"
BASE COUNT 427 a 785 c 821 g 394 t
ORIGIN
Query Match 3.7%: Score 118.6; DB 9; Length 2427;
Best Local Similarity 56.1%: Pred. No. 2.2e-12;
Matches 270; Conservative 0; Mismatches 199; Indels 12; Gaps 2;
QY 421 ccgagcttcacacagcctccaagatgagcgccggtgatcgacgagcccggtgagtc 480
Db 451 CCTTACTGACACGGCCCGAGCGGATGGACAAGAGTGTCTGGCTGCGCCGCGCCAC 510
QY 481 tccgtgagctcaagtgtggtgagcgagccctcgccagcatcaagtgtgagaa 540
Db 511 ACCGTCCGCTTCCGCTGCGACCGCCTGTGACACCCCACTCCCTCATCTCTGCTGAAG 570
QY 541 gacgacagcgctt-----gacgagccagagcgccgtgtgagccaggaagaag 591
Db 571 AACGGAGGAGGATTCGCGCGGAGACCGCATTTGGAGGATCAAGCTGGGCATCAGCAG 630
QY 592 tggacaacttgagccttgaagaactctgcccggagagcagcggaataaactcgccgctg 651
Db 631 TGGAGCTGTGATGTAAGAAAGCTGTGTGCTTGGACCGGGCACTTACACTGCTGCTG 690

```

[illegible]

RESULT	39			
LOCUS	AX047605			
DEFINITION	Sequence 10 from Patent WO0068424.	PAT	15-DEC-2000	
ACCESSION	AX047605			
VERSION	AX047605.1			
KEYWORDS	GI:11876673			
SOURCE	synthetic construct.			
ORGANISM	synthetic construct.			
REFERENCE	artificial sequence.			
AUTHORS	1 (bases 1 to 2427)			
TITLE	Cappellen, D., Chopin, D., Radvanyi, F., Ricol, D. and Thierry, J.P.			
JOURNAL	Means for detecting and treating pathologies linked to fgfr3			
	Patent: WO 0068424-A 10 16-NOV-2000;			
	Institut Curie (FR) ; CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE			
	(CNRS) (FR)			

FEATURES	Location/Qualifiers
SOURCE	1. 2427 /organism="synthetic_construct" /db_xref="taxon:32630" /note="Mutant X809G GGFRA3-IIIB: (Mutant 1)"
BASE COUNT	428 a 784 c 822 g 393 t
ORIGIN	
Query Match	3.7%: Score 118.6; DB 9; Length 2427;
Best Local Similarity	56.1%: Pred. No. 2.2e-12;
Matches 270; Conservative	0; Mismatches 199; Indels 12; Gaps 2

	Query Match	3.7%	Score 118.6;	DB 9;	Length 2427;
	Best Local Similarity	56.1%;	Pred. No. 2.2e-12;		
	Matches 270;	Conservative	0;	Mismatches 199;	Indels 12; Gaps
0y	421	ccgcgcctcacacagcgcctcccaagatgagcgccggttgatcgcagcgcgcgtggtgagc	480		
Db	451	CTTACTGTGGACACGGCCCGACGGATGGACAAGAGCTGCTGGCGCTGGCGGCCAAC	510		
0y	481	tcctgtcgcgcccaagtgtcggtggccagcggcgacccctcgcccgccatccatcgttgatgaag	540		
Db	511	ACCGTCCGCTTCGCTGCGCCACGGCTGGCAACCCCACTCCCTTCATCTCTGTGGCTGAAG	570		
0y	541	gacgacccagcgcctt-----gacgcgcgcacagcgccgtgagcccgccagaaagaag	591		
Db	571	AACGGCAGAGGAGTTCCGCGCGGACACCGCATTTGGAGGACATCAAGCTGCGGCATACAGAG	630		
0y	592	tggaacactgaagccttgaaagacacttcgcgcgcggagagacagcgcgcaataaacactgcgcgtg	651		
Db	631	TGGACCTGTGTCATGGAGAAAGCGTGTGTCCTTCGGACCGCGGCACACTACACTGCTGTG	690		
0y	652	tcgaacgcgcgcgcgcgcacataacgcgcacactacaagttgtagttgatccagcgcgaacct	711		
Db	691	GAGAACACAGTTTGGCGACGATCCGCGACAGCACTACGCTGGACGTCGTGAGAGCGTCCCG	750		
0y	712	tcgaagcccgctgctccacagcagcaccgcgttgaaacacagacggttggaacttcggggggacc	771		

Db	751	CACGCGCCCATCTCTGCAGAGCGGGGCTGCGCGGCCNAACCAAGAGGGGGTGTGGGGCAGGAC	810
Qy	772	acgtctctccagtgcaagatgctgcagcgagcgaatgaagccgylatccagtgagctgaagcgc	831
Db	811	GTGGAGTTCCACTGCAAGAGTGTACAGTGCAGCAGACGCCCCACATCCAGTGGCTCAACAC	870
Qy	832	gtgga-----gtacgagcccgagagggccgcagcaactccacatcgatgtgtggcgccgcaag	888
Db	871	GTGGAGGTGAACGGCAGAGAGGTGGGGCCCGACGCAACACCTCTACCTTACCGTGCTCAAG	930
Qy	889	t	889
Db	931	T	931

RESULT	40
AX047606	
LOCUS	AX047606 2427 bp DNA PAT 15-DEC-2000
DEFINITION	Sequence 11 from Patent WO0068424.
ACCESSION	AX047606
VERSION	AX047606.1 GI:11876674
KEYWORDS	.
SOURCE	synthetic construct.
ORGANISM	synthetic construct.
REFERENCE	artificial sequence.
AUTHORS	1 (bases 1 to 2427)
TITLE	Cappellen,D., Chopin,D., Radvanyi,F., Ricol,D. and Thery,J.P.
JOURNAL	Means for detecting and treating pathologies linked to fgfr3 Patent: WO 0068424-A 11 15-NOV-2000; Institut Curie (FR) ; CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR)

FEATURES	Location/Qualifiers
source	1..2427
	/organism="synthetic construct" /db_xref="taxon:32630" /note="Mutant X809G FGFR3-IIIb: (Mutant 2)"
BASE COUNT	429 a 784 c 821 g 393 t
ORIGIN	
Query Match	3.7%: Score 118.6: DB 9: length 2427;
Best Local Similarity	56.1%: Pred. No. 2.2e-12;
Matches 270; Conservative	0; Mismatches 199; Indels 12; Gaps 2

Query Match	3.7%;	Score 118.6;	DB: 9;	Length 2427;
Best Local Similarity	56.1%;	Pred. No. 2.2e-12;		
Matches 270;	Conservative	0;	Mismatches 199;	Indels 12; Gaps
OY 421	ccgcgttcacagccgcctcccaagatgagcgccgggtgatcgcagcgccctggtgtagc	480		
Db 451	ccttactctgacacacggcccgacggcgatgatgacaaagaaactgctggccgctggccggccac	510		
OY 481	tcctgtgcggtcctaagatgcgtgtggtccagcggtgacacccctcgtgcccagatcactgtgatgaag	540		
Db 511	aaccgtcccttccgtccgctgcgccacggcgacacccacacccctccctccacatcctctgacatgaag	570		
OY 541	gacgcacagccctt-----gacgcgcccaagcgctgtgacccagaaagaaag	591		
Db 571	aacggcagaggaattccggcgacacaccgattttggagagcatcaagctggcggcatcacag	630		
OY 552	tggacactgagccttgaagaacctctgcgcggaggaagacagcgggcaaatacacctctgcgcgtg	651		
Db 631	tggacacctggtcatctggaagagctgctggctcccgacccggcgaactctgctgctgctg	690		
OY 652	tcgaacgcgcggcgccatctaaagcgaacctctacaagttgatctgacgcagcggaacctgt	711		
Db 691	gagaaacaggtttggcagacatccggcagacatgacacgctggacgtgcttggaagcgtctcccg	750		
OY 712	tccaagcccgctgtctacagcgacgcaccccggtgaacaagcagctgtgaccttcgggggagcc	771		
Db 751	caccggcccatctctgcagagccggcgctgcggcccaaccagagcgcggtctggcgacgac	810		
OY 772	acgtctctcagatgacaggttgcagcgagccttgagcccgctgtgacagctgtgtcgaagcgc	831		
Db 811	gtgcagtttccactgcacagcttacaagtacccacacgcccacatctcagctgacacac	870		
OY 832	gttga-----gtacgcgcgcgcgaaggtccgcacaaacttcacacatctgattgtggcggtcgaag	888		

Db 871 GTGAGGTGAACGCGACGAGGTGGCCCGGACGCGACACCTTACGTTACGCTGCTCAAG 930  
 Oy 889 t 889  
 Db 931 T 931

RESULT 41  
 AX047607 2427 bp DNA PAT 15-DEC-2000

LOCUS AX047607  
 DEFINITION Sequence 12 from Patent WO0068424.  
 ACCESSION AX047607  
 VERSION AX047607.1 GI:11876675

KEYWORDS  
 SOURCE synthetic construct.  
 ORGANISM synthetic construct.

REFERENCE 1 (bases 1 to 2427)  
 AUTHORS Cappellen,D., Chopin,D., Radvanyi,F., Ricol,D. and Thierry,J.P.  
 TITLE Means for detecting and treating pathologies linked to Igfr3  
 JOURNAL Patent: WO 0068424-A 12 16-NOV-2000;  
 Institut Curie (FR) ; CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE  
 (CNRS) (FR)

# FEATURES

source Location/Qualifiers  
 1..2427  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 /note="Mutant X809g FGFR3-IIID: (Mutant 3)"

BASE COUNT 428 a 785 c 821 g 393 t  
 ORIGIN  
 Query Match 3.7%; Score 118.6; DB 9; Length 2427;  
 Best Local Similarity 56.1%; Pred. No. 2.2e-12;  
 Matches 270; Conservative 0; Mismatches 199; Indels 12; Gaps 2;

Oy 421 ccgcgccttcacagccctccaagaatgagcgcgcgtgtagtcacagcccggtggtagc 480  
 Db 451 CCTTACTGAGACGCGCCGAGCGGAGTGAACAAGCTGCTGGCCGCGCGGCCGCAAC 510  
 Oy 481 tccgtgagctcaagtgtcgtgcaagcgagccctcgcccgacatcacgttgatgaag 540  
 Db 511 ACGGTCCGCTTCGCTCCGACCGCGTGGCAACCCCATCTCTCTGCTGGAAG 570  
 Oy 541 gaagaccagcctt-----gaagcgcccaagagcgctgtagcccaagaagaag 591  
 Db 571 AACGGCAGGGGATTCCGCGGAGACACCGATTGAGGATCAAGCTGCGGATCAAGCAG 630  
 Oy 592 tggacactgagcctgtagaagacctgcgccgagagacagcgcaaatcacctgacgctg 651  
 Db 631 TGGAGCTTGATGATGAAGAGCTGTGCTCGACGCGGCAACTACACTGGTGTG 690  
 Oy 652 tcgaacgcgcgcgcgcacatcaagcgcacactacaagaatgtagtgcagcgagaccgt 711  
 Db 691 GAGAACAGATTGTGACAGATCCGAGACGATACAGCTGAGCGTGTGAGCGGTCCG 750  
 Oy 712 tcaagcccggtgtcacaagcagcaccgctgtagcagcggtgtagcttcgagggagc 771  
 Db 751 CACCGGCCATCTGACAGCGCGGCTGCGGCCAACACAGAGGCGGTGCTGGGACGAC 810  
 Oy 772 agctcttcagtcgaagtgtcagcagcagtgtagcagtgtagctgagtcgagcgc 831  
 Db 811 GTGAGATTCCACTGCAAGGTGTACAGTGACGACACCCCAATCCAGTGGCTCAAGCAG 870  
 Oy 832 gtgga---gtacgcgcgcgcgcgcgcacacacacacacacacacacacacacacac 888  
 Db 871 GTGAGGTGAACGCGACGAGGTGGCCCGGACGCGACACCTTACGTTACGCTGCTCAAG 930  
 Oy 889 t 889  
 Db 931 T 931

RESULT 42  
 AX047608 2427 bp DNA PAT 15-DEC-2000

LOCUS AX047608  
 DEFINITION Sequence 13 from Patent WO0068424.  
 ACCESSION AX047608  
 VERSION AX047608.1 GI:11876676

KEYWORDS  
 SOURCE synthetic construct.  
 ORGANISM synthetic construct.

REFERENCE 1 (bases 1 to 2427)  
 AUTHORS Cappellen,D., Chopin,D., Radvanyi,F., Ricol,D. and Thierry,J.P.  
 TITLE Means for detecting and treating pathologies linked to Igfr3  
 JOURNAL Patent: WO 0068424-A 13 16-NOV-2000;  
 Institut Curie (FR) ; CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE  
 (CNRS) (FR)

# FEATURES

source Location/Qualifiers  
 1..2427  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 /note="Mutant X809L FGFR3-IIID:"

BASE COUNT 428 a 784 c 820 g 395 t  
 ORIGIN  
 Query Match 3.7%; Score 118.6; DB 9; Length 2427;  
 Best Local Similarity 56.1%; Pred. No. 2.2e-12;  
 Matches 270; Conservative 0; Mismatches 199; Indels 12; Gaps 2;

Oy 421 ccgcgccttcacagccctccaagaatgagcgcgcgtgtagtcacagcccggtggtagc 480  
 Db 451 CCTTACTGAGACGCGCCGAGCGGAGTGAACAAGCTGCTGGCCGCGCGGCCGCAAC 510  
 Oy 481 tccgtgagctcaagtgtcgtgcaagcgagccctcgcccgacatcacgttgatgaag 540  
 Db 511 ACGGTCCGCTTCGCTCCGACCGCGTGGCAACCCCATCTCTCTGCTGGAAG 570  
 Oy 541 gaagaccagcctt-----gaagcgcccaagagcgctgtagcccaagaagaag 591  
 Db 571 AACGGCAGGGGATTCCGCGGAGACACCGATTGAGGATCAAGCTGCGGATCAAGCAG 630  
 Oy 592 tggacactgagcctgtagaagacctgcgccgagagacagcgcaaatcacctgacgctg 651  
 Db 631 TGGAGCTTGATGATGAAGAGCTGTGCTCGACGCGGCAACTACACTGGTGTG 690  
 Oy 652 tcgaacgcgcgcgcgcacatcaagcgcacactacaagaatgtagtgcagcgagaccgt 711  
 Db 691 GAGAACAGATTGTGACAGATCCGAGACGATACAGCTGAGCGTGTGAGCGGTCCG 750  
 Oy 712 tcaagcccggtgtcacaagcagcaccgctgtagcagcggtgtagcttcgagggagc 771  
 Db 751 CACCGGCCATCTGACAGCGCGGCTGCGGCCAACACAGAGGCGGTGCTGGGACGAC 810  
 Oy 772 agctcttcagtcgaagtgtcagcagcagtgtagcagtgtagctgagtcgagcgc 831  
 Db 811 GTGAGATTCCACTGCAAGGTGTACAGTGACGACACCCCAATCCAGTGGCTCAAGCAG 870  
 Oy 832 gtgga---gtacgcgcgcgcgcgcgcacacacacacacacacacacacacacacac 888  
 Db 871 GTGAGGTGAACGCGACGAGGTGGCCCGGACGCGACACCTTACGTTACGCTGCTCAAG 930  
 Oy 889 t 889  
 Db 931 T 931

## RESULT 43

AX047609 2427 bp DNA PAT 15-DEC-2000  
 LOCUS AX047609  
 DEFINITION Sequence 14 from Patent WO0068424.  
 ACCESSION AX047609  
 VERSION AX047609.1 GI:11876677



KEYWORDS	synthetic construct.
SOURCE	synthetic construct.
ORGANISM	synthetic construct.
REFERENCE	artificial sequence.
AUTHORS	1 (bases 1 to 2427)
TITLE	Cappelletten, D., Choplin, D., Radvanly, F., Ricoul, D. and Phier, J. P.
JOURNAL	Means for detecting and treating pathologies linked to fgfr3
	Patent: WO 0068424-A 14 16-NOV-2000;
	Institut Curie (FR) ; CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
	(CNRS) (FR)
FEATURES	Location/Qualifiers
source	1..2427
	/organism="synthetic construct"
	/db_xref="taxon:32630"
	/note="Mutant N542K FGFR3-IIIB: (Mutant 1)"
BASE COUNT	429 a 783 c 821 g 394 t
ORIGIN	
Query Match	3.7%; Score 118.6; DB 9; Length 2427;
Best Local Similarity	56.1%; Pred. No. 2.2e-12;
Matches 270; Conservative	0; Mismatches 199; Indels 12; Gaps 2.
OY	421 ccgagcttcacacagccctcccaagatgtagggcgaggtatgcacagggccgtggtagc 480
Db	451 CCTTACTGAGACCGGCGCCGAGCGGATGGACAAAGACTGCTGGCCGCGCCGCAAC 510
OY	481 tccgtgcggtctaaagtgcgttgccagcgagccacctgcgcccagatcacgttgatgaag 540
Db	511 ACCGTCGCTTCGCGTGGCCAGCCGCTGGCAACCCACCTCCATCTCCGCTGGCGAAG 570
OY	541 gaagcagcagcctt-----gaagcgccagagcgcgctgtagcccaagaagaag 591
Db	571 AACGGCAGGAGATTCCGGCGGAGCAGCACCGCATTTGAGAGGCATCAAGCTCGGCATCAGCAG 630
OY	592 tggagacgtgagcgtctaaagacctgagcccgagagcagagaggaataacacctgcgcgag 651
Db	631 TGGACCTGCTCATGGAAGCGCTGTGCTTCGGACCGGGCACTACACCTGCGTCGTG 690
OY	652 tcgaacgcgcgcgcgcgcacacacgcacacaaagtgtgatgtatccagcgagaccgt 711
Db	691 GAGAACAGATTGGCAGAGATCCGGCAGACGATACAGCTGTGAGTGTGAGAGCGCTCCCG 750
OY	712 tccaaagccgcgtctacagcagcagcaccgccgtgaacaagaaggttgagcttcgaggagacc 771
Db	751 CACCGCCCATCTCTCAGAGCGGCGCTGGCCGCAACAGACGCGGCTGTGGGCAAGCAG 810
OY	772 agctctctcaggtgtaaggttgccagcgagcgtgaaagccggtgatctcagtggtctgaagcgc 831
Db	811 GTGAGTTTCACATGCAAGGTATGACAGTACGACAGCCCACTCAGAGGCTCAAGCAG 870
OY	832 gtgcga---gtacgcgcgcgcgcgcgcacacacacacacacacatcgatgtggcgccagag 888
Db	871 GTGAGGAGTGAACGCGACGACAGGTGGGCCCGGACGCGCACACCTTACGTTACCTGCTCAAG 930
OY	889 t 889
Db	931 t 931
RESULT 44	
LOCUS	AX047610 2427 bp DNA PAT 15-DEC-2000
DEFINITION	Sequence 15 from Patent WO0068424.
ACCESSION	AX047610
VERSION	AX047610.1 GI:11876678
KEYWORDS	
SOURCE	synthetic construct.
ORGANISM	synthetic construct
REFERENCE	artificial sequence.
AUTHORS	1 (bases 1 to 2427)
TITLE	Cappelletten, D., Choplin, D., Radvanly, F., Ricoul, D. and Phier, J. P.
	Means for detecting and treating pathologies linked to fgfr3
	Patent: WO 0068424-A 14 16-NOV-2000;
	Institut Curie (FR) ; CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
	(CNRS) (FR)

[illegible]

BASE COUNT                    /note="Mutant G382R FGFR3-IIIB: (Mutant 1)"  
429 a                    784 c                    820 g                    394 t  
ORIGIN

Query Match                    3 7%;    Score 118.6;    DB 9;    Length 2427;  
Best Local Similarity    56.1%;    Pred. No. 2.2e-12;  
Matches 270;    Conservative    0;    Mismatches 199;    Indels 12;    Gaps 2;

QY    421 ccgcgtccacacagccctccacaagatgaagcgccggtgatcgaagcgcgtggtatgc 480  
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
DB    451 CCTTACTGACACAGCGGCCGAGATGACACAGAGCTCTGCGCCGCGGCCGCCAAC 510  
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY    481 tccgtgcggtcaagtgcgtgagccagcgagccctcgcccgacatcaagtgtgataag 540  
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
DB    511 ACCGTCGCTTCGCTGCCACAGCCGCTGGCAACCCACATCCTCCATCTCTGCTGAAG 570  
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY    541 gacgacacagcctt-----gacgagcccaagagccgctgagcccaagagaagaag 591  
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
DB    571 AACGCGAGGAGTCCGCGCGAGCACCGCATTTGGAGGCATCAAGCTGGCGCATCAGCAG 630  
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY    592 tggacactgagcctgaagaacctgcccggaggaagcggcaatatacactgcgcgtg 651  
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
DB    631 TGGAGCTGTGATGGAAGCGTGTGCTCGACCGCGCACTACACTGCGTGTG 690  
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY    652 tcgaaccgcgagggcgccatcaagccacctacaaggtgatagtatccagcggaccgt 711  
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
DB    691 GAGAACACAGTTTGGCAGCATCCGCGACAGCATACGCTGACGTGAGCGCTGCCCG 750  
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY    712 tcgaagcccggtgtcaacagcaaccccgtaacacagcgtgtgacttcggggagcc 771  
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
DB    751 CACCGGCCCATCTCTGACAGCGGGGCTGCCGCGCAACAGAGCGGCTGTGGGACAGCAG 810  
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY    772 acgtccttcagtgcaaggtgcagcgaagcgtgagcgggtgatccagtgtgctgaagcg 831  
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
DB    811 GTGAGTTCACACTCAAGGTAAGTACAGTACAGCAGCCACATCCAGTGGCTCAAGCAG 870  
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY    832 gtgag---gtacggcgcgagggcgcccaaacctccacatcgatgtggcgcgcaagaag 888  
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
DB    871 GTGAGGTGAACGAGACAGAGTGGGCCGAGACGCAACCTTACCTTACCTGCTCAAG 930  
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
QY    889 t 889  
      |  
DB    931 T 931

Search completed: August 17, 2001, 05:01:58  
Job time: 9534 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2001, 02:22:14 ; Search time 2317.92 Seconds

(without alignments)  
12993.024 Million cell updates/sec

Title: US-09-598-042A-2

Perfect score: 3186  
Sequence: 1 ggcgaattggccgcgaacgtcg.....tttccctcccaaaaaaaaaa 3186

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estda:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*
44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estp11:*
73: em_estp12:*
74: em_estp13:*
75: em_estp14:*
76: em_estp15:*
77: em_estp16:*
78: em_estp17:*
79: em_estp18:*
80: em_estp19:*
81: em_estp110:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*
```

117: gb\_est48:\*  
118: gb\_est49:\*  
119: gb\_est50:\*  
120: gb\_est51:\*  
121: gb\_est52:\*  
122: gb\_est53:\*  
123: gb\_est54:\*  
124: gb\_est55:\*  
125: gb\_est56:\*  
126: gb\_est57:\*  
127: gb\_est58:\*  
128: gb\_est59:\*  
129: gb\_est60:\*  
130: gb\_est61:\*  
131: gb\_est62:\*  
132: gb\_est63:\*  
133: gb\_est64:\*  
134: gb\_est65:\*  
135: gb\_est66:\*  
136: gb\_est67:\*  
137: gb\_est68:\*  
138: gb\_est69:\*  
139: gb\_est70:\*  
140: gb\_est71:\*  
141: gb\_est72:\*  
142: gb\_est73:\*  
143: gb\_est74:\*  
144: gb\_est75:\*  
145: gb\_est76:\*  
146: gb\_est77:\*  
147: gb\_est78:\*  
148: gb\_est79:\*  
149: gb\_est80:\*  
150: gb\_est81:\*  
151: gb\_est82:\*  
152: gb\_est83:\*  
153: gb\_est84:\*  
154: gb\_est85:\*  
155: gb\_est86:\*  
156: gb\_est87:\*  
157: gb\_est88:\*  
158: gb\_est89:\*  
159: gb\_est90:\*  
160: gb\_est91:\*  
161: gb\_est92:\*  
162: gb\_est93:\*  
163: gb\_est94:\*  
164: gb\_est95:\*  
165: gb\_est96:\*  
166: gb\_est97:\*  
167: gb\_est98:\*  
168: gb\_est99:\*  
169: gb\_est100:\*  
170: gb\_est101:\*  
171: gb\_est102:\*  
172: gb\_est103:\*  
173: gb\_est104:\*  
174: gb\_est105:\*  
175: gb\_est106:\*  
176: gb\_est107:\*  
177: gb\_est108:\*  
178: gb\_est109:\*  
179: gb\_est110:\*  
180: gb\_est111:\*  
181: gb\_est112:\*  
182: gb\_est113:\*  
183: gb\_est114:\*  
184: gb\_est115:\*  
185: gb\_est116:\*  
186: gb\_est117:\*  
187: gb\_est118:\*  
188: gb\_est119:\*  
189: gb\_est120:\*

190: gb\_est121:\*  
191: gb\_est122:\*  
192: gb\_est123:\*  
193: gb\_est124:\*  
194: gb\_est125:\*  
195: gb\_est126:\*  
196: gb\_est127:\*  
197: gb\_est128:\*  
198: gb\_est129:\*  
199: gb\_est130:\*  
200: gb\_est131:\*  
201: gb\_est132:\*  
202: gb\_est133:\*  
203: gb\_est134:\*  
204: gb\_est135:\*  
205: gb\_est136:\*  
206: gb\_est137:\*  
207: gb\_est138:\*  
208: gb\_est139:\*  
209: gb\_est140:\*  
210: gb\_est141:\*  
211: gb\_est142:\*  
212: gb\_est143:\*  
213: gb\_est144:\*  
214: gb\_est145:\*  
215: gb\_est146:\*  
216: gb\_est147:\*  
217: gb\_est148:\*  
218: gb\_est149:\*  
219: gb\_est150:\*  
220: gb\_est151:\*  
221: gb\_est152:\*  
222: gb\_est153:\*  
223: gb\_est154:\*  
224: gb\_est155:\*  
225: gb\_est156:\*  
226: gb\_est157:\*  
227: gb\_est158:\*  
228: gb\_est159:\*  
229: gb\_est160:\*  
230: gb\_est161:\*  
231: gb\_est162:\*  
232: gb\_est163:\*  
233: gb\_est164:\*  
234: gb\_est165:\*  
235: gb\_est166:\*  
236: gb\_est167:\*  
237: gb\_est168:\*  
238: gb\_est169:\*  
239: gb\_est170:\*  
240: gb\_est171:\*  
241: gb\_est172:\*  
242: gb\_est173:\*  
243: gb\_est174:\*  
244: gb\_est175:\*  
245: gb\_est176:\*  
246: gb\_est177:\*  
247: gb\_est178:\*  
248: gb\_est179:\*  
249: gb\_est180:\*  
250: gb\_est181:\*  
251: gb\_est182:\*  
252: gb\_est183:\*  
253: gb\_est184:\*  
254: gb\_est185:\*  
255: gb\_est186:\*  
256: gb\_est187:\*  
257: gb\_est188:\*  
258: gb\_est189:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	719.6	22.6	978	147	BF305985	BF305985 601889282
2	694.8	21.8	910	139	BE746024	BE746024 601578548
3	675.4	21.2	854	107	AU124593	AU124593 601524593
4	674	21.2	685	115	AM411524	AM411524 fh12911.x
5	661.2	20.8	912	169	BF792014	BF792014 602252413
6	655.6	20.6	819	32	AV684577	AV684577 AV684577
7	651.8	20.5	697	115	AM411525	AM411525 fh12911.y
8	646	20.3	829	140	BE793458	BE793458 601588524
9	644.2	20.2	694	165	BE272083	BE272083 601140878
10	643.6	20.2	1224	139	BE745264	BE745264 601574033
11	634.4	19.9	687	153	BG423273	BG423273 602450596
12	631.8	19.8	876	172	BG036629	BG036629 602326830
13	632	19.5	983	152	BG331236	BG331236 602431952
14	598.6	18.8	753	166	BE378421	BE378421 60136755
15	597.4	18.8	815	171	BF968051	BF968051 602269220
16	595.6	18.7	727	152	BG326224	BG326224 602424936
17	594.8	18.7	735	152	BG32879	BG32879 602430724
18	589.2	18.5	957	153	BG423479	BG423479 602449337
19	586.2	18.4	668	165	BE273980	BE273980 601104454
20	570.4	17.9	837	143	BF033645	BF033645 60153948
21	565.4	17.7	1122	172	BF980548	BF980548 602304311
22	564.6	17.7	974	147	BF340012	BF340012 602036948
23	558	17.5	945	153	BG422503	BG422503 602449078
24	552	17.3	913	141	BE873767	BE873767 601483916
25	543.2	17.0	984	172	BF970020	BF970020 602272864
26	535.2	16.8	936	172	BF984116	BF984116 602308605
27	529	16.6	529	148	BF436316	BF436316 7p06b06.x
28	521.6	16.4	660	121	AM645240	AM645240 QV0-C7001
29	519.4	16.3	873	147	BF306936	BF306936 601891451
30	514	16.1	574	108	AU148420	AU148420 AU148420
31	510.2	16.0	774	175	BG254594	BG254594 602368534
32	501.4	15.7	762	171	BF968611	BF968611 602271034
33	500	15.7	829	141	BE898179	BE898179 601435523
34	494.6	15.5	961	147	BF306438	BF306438 601893323
35	493.4	15.5	672	140	BE77873	BE77873 601463636
36	489	15.3	537	32	AV707425	AV707425 AV707425
37	465	14.6	604	150	BF585002	BF585002 602098902
38	459.2	14.4	981	147	BF307860	BF307860 601890670
39	455.4	14.3	874	153	BG387406	BG387406 602456172
40	453.6	14.2	951	140	BE780737	BE780737 601469609
41	444	13.9	456	103	A1921390	A1921390 wo24c03.x
42	443	13.9	443	111	AM083114	AM083114 xc06b10.x
43	442	13.9	487	20	A1432643	A1432643 t122801.x
44	437	13.7	944	172	BF969458	BF969458 602271754
45	436.2	13.7	441	7	AA442297	AA442297 zv54c08.r

ALIGNMENTS

RESULT	1	798 bp	mrna	EST	21-NOV-2000
LOCUS	BF305985				
DEFINITION	601889282f1 NIH_MGC_17	Homo sapiens	cdna clone	IMAGE:4122981.5'	
ACCESSION	BF305985				
VERSION	BF305985.1	GI:11253033			
KEYWORDS	EST				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	NIH-MGC http://mgc.nci.nih.gov/				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
COMMENT	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				

Email: cgaabs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov  
Plate: LNCM1007 row: d column: 22  
High quality sequence stop: 658.  
Location/Qualifiers  
1. 978

FEATURES

source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NIH\_MGC\_17"  
/clone\_lib="NIH\_MGC\_17"  
/tissue\_type="rhodomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: muscle; Vector: pORF7; Site:1: EcoRI; Site:2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGGAG(g). Size-selected for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."  
BASE COUNT 292 a 299 c 226 g 161 t  
ORIGIN

Query Match 22.6%; Score 719.6; DB 147; Length 978;  
Best Local Similarity 95.3%; Pred. No. 9.4e-155;  
Matches 785; Conservative 0; Mismatches 34; Indels 5; Gaps 4;

Qy	1897	cacagaaatgctgcgaatgacacgacacgacagagagatgacagacatataaagac	1956
Db	2	CACAGAAAGTGTGCCAAATGSCACGACGACAGAGACATGCCAACAATACAAAGAC	61
Qy	1957	atgctgctgaacatacacaacgacacccatgacagatgctgctgacacacac	2016
Db	62	ATGCTGCTGTGAATATACACGACACACCATGCTGCTGCTGACACACAC	121
Qy	2017	acaacgagatgctgctgacgacacacgctgacagatgctgacacacacgt	2076
Db	122	ACACACGATATGCTGTGACGACACACGTCAGATATGATCGGACACACG	181
Qy	2077	gcacagatgctgctgacacacacagataatgctgctgacacacacatgacagata	2136
Db	182	GCACAGATATGCTGCTGACACACAGATATGCTGCTGACACACATGCGGATA	241
Qy	2137	ctgctgacacacacacacacacgctgacagatgctgctgacacacacacacac	2196
Db	242	TTGCTGACACACACACACACACGTCGACAGATATGCTGTGACACACAC	301
Qy	2197	ttgcagatgctgctgacacacacacttccagacacacgctgacagcgcagatgctg	2256
Db	302	TTCAGATATGCTGCTGACACACACCTTCAGACACACGTCACAGCGCAGATATC	361
Qy	2257	cctgacacacgagatgctgctgacacacacacacacacgacacacacacacacac	2316
Db	362	CTTGACACACGAGATATGCTGTAGTCACACACACACACACATGCTGCGGACA	421
Qy	2317	cacacacgacatgacagatgctgctgacacacacacacacacgacacacacacac	2376
Db	422	CACACACGATGACAGATATGCTGCTGACACACACACACACGAGATATGCTG	481
Qy	2377	gacacacacacagatgctgctgacacacacacacacacacacacacacacacacac	2435
Db	482	GACACACACACAGATATGCTGACTCAAAACATCACACACGTCAGATATGCTG	541
Qy	2436	cacacagatgacagatgctgctgacacacacacacacacacacacacacacacacac	2495
Db	542	CACACATGTCACAGATATGCTGTGACATGACACACACACATGATATGCTGCGGA	601

[illegible]



Matches	677:	Conservative	0:	Mismatches	5:	Indels	0:	Gaps	0:
QY 1674	actggagagatggaagacgagctgcagacgaaggcaggagcccatgycgagaggaat	1733							
DB 4	acgagagagatgagagacgagctgcagacgaaggcaggagcccatgycgagaggaat	63							
QY 1734	ggccagcaaccccaagcagctgtgtgtgtgagatagccccctggacacacacacagaca	1793							
DB 64	ggccagcaaccccaagcagctgtgtgtgtgagatagccccctggacacacacacagaca	123							
QY 1794	caacacacccctggatgcatgtatgcacacacatgcgcgcacagctgtccctggaagga	1853							
DB 124	caacacacccctggatgcatgtatgcacacacatgcgcgcacagctgtccctggaagga	183							
QY 1854	caagtcagcaacacacagcagatgcagacagatagcgcgcctggcagacagatagctccca	1913							
DB 184	caagtcagcaacacacacagcagatgcagacagatagcgcgcctggcagacagatagctccca	243							
QY 1914	aatgcagcagacacacacagcagatgcagacacacacacacacacacacacacacacac	1973							
DB 244	aatgcagcagacacacacacagcagatgcagacacacacacacacacacacacacacac	303							
QY 1974	acacgac	2033							
DB 304	acacgac	363							
QY 2034	ctggagcagac	2093							
DB 364	ctggagcagac	423							
QY 2094	ggac	2153							
DB 424	ggac	483							
QY 2154	caac	2213							
DB 484	caac	543							
QY 2214	gaac	2273							
DB 544	gaac	603							
QY 2274	atgctgtctatgcac	2333							
DB 604	atgctgtctatgcac	663							
QY 2334	atatgtcttcggaac	2355							
DB 664	atatgtcttcggaac	685							
RESULT 5									
LOCUS	BF792014	912 bp	EST	12-JAN-2001					
DEFINITION	60223243f1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4344674 5'								
ACCESSION	BF792014								
VERSION	BF792014.1	GI:12097068							
KEYWORDS	EST.								
SOURCE	human.								
ORGANISM	Homo sapiens								
REFERENCE	National Institutes of Health, Mammalian Gene Collection (MGC)								
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/								
TITLE	Unpublished (1999)								
JOURNAL	Contact: Robert Strausberg, Ph.D.								
COMMENT	Email: cgaabs-remail.nih.gov								
	Tissue Procurement: ATCC								
	CDNA Library Preparation: Life Technologies, Inc.								
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)								
	DNA Sequencing by: Incyte Genomics, Inc.								

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov  
Plate: LHM963 row: 3 column: 03  
High quality sequence stop: 647.  
Location/Qualifiers

#### FEATURES

source

1. .912  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4344674"  
/issue\_type="adrenal cortex carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: adrenal gland; Vector: pcMV-SPORE6; Site: 1; NotI; Site: 2; SalI; Cloned unidirectionally; Oligo-dT primed. Average insert size 1.229 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

BASE COUNT 228 a 280 c 237 g 167 t  
ORIGIN

Query Match 20.8%; Score 661.2; DB 169; Length 912;  
Best Local Similarity 95.0%; Pred. NO. 2.4e-141;  
Matches 737; Conservative 0; Mismatches 33; Indels 6; Gaps 5;

QY 1584	gtggagggcagaggtccaccagacatccatcatcagctgtagagcgacacgtatctgca	1643							
DB 1	gtggagggcagaggtccaccagacatccatcatcagctgtagagcgacacgtatctgca	60							
QY 1644	agggcagggggggcggcgacacagcagacacacacacacacacacacacacacacacac	1703							
DB 61	tgggcagcggggggggcggcgacacacacacacacacacacacacacacacacacacac	120							
QY 1704	gaagcagagggggac	1763							
DB 121	gaagcagagggggac	180							
QY 1764	gatatgccccctggac	1823							
DB 181	gatatgccccctggac	240							
QY 1824	acatgctgcgac	1883							
DB 241	acatgctgcgac	300							
QY 1884	tgcgccttggcagac	1943							
DB 301	tgcgccttggcagac	360							
QY 1944	aacatacaagac	2003							
DB 361	aacatacaagac	420							
QY 2004	tggac	2063							
DB 421	tggac	480							
QY 2064	cggac	2123							
DB 481	cggac	540							
QY 2124	acatgacagac	2183							
DB 541	acatgacagac	600							
QY 2184	acacgac	2242							
DB 601	acacgac	660							
QY 2243	ggcagac	2298							
DB 661	ggcagac	720							



Oy	2299	agaactgctcgt-cggagacacacacgcgcgtcacgatatgcttcgccgacacaa	2253
Dd	721	CAGAACTGCTGCCGGACACACACGCGCCTTCCACAGATTTCGTTCACACACA	776
RESULT	6		
LOCUS	AV684577	819 bp	mRNA
DEFINITION	AV684577 GKC Homo sapiens CDNA clone GKCDP01 5', mRNA sequence.	EST	25-SEP-2000
VERSION	AV684577.1 GI:10286440		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 819)		
AUTHORS	Mu,T., Qian,B., Huang,Y., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H. Xu,X., Li,N., Peng,Q., Liu,F., Ou,J., Song,H., Cheng,Z., Zeng,L., XU,S.S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang, Y., Gu,Y., Chen,Z. and Han,Z. Homo sapiens CDNA GK-clones Unpublished (2000) Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel.: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai.		
TITLE	JOURNAL		
COMMENT			
FEATURES	Source		
	Location/Qualifiers		
	1..819	/organism="Homo sapiens"	
	/db_xref="taxon:9606"		
	/clone="GKCDP01"		
	/clone_lib="GKC"		
	/tissue_type="hepatocellular carcinoma"		
	/dev_stage="Adult"		
	/lab_host="SOLR"		
	/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"		
BASE COUNT	230 a 253 c 220 g 115 t	1 others	
ORIGIN			
Query Match	20.6%; Score 655.6; DB 32; Length 819;		
Best Local Similarity	92.8%; Pred. NO. 4.6e-140;		
Matches 743; Conservative	0; Mismatches 50; Indels 8; Gaps 5		
Oy	1545	cacaacacacacacacacacacctccacacacactcaacacactcaacacgctgtaggaaggcaccag	1604
Dd	8	CAGACACGGGCGCACACACTCTCACACACTCATACAGCTGGAGGGCGCAGGTCCACCAT	67
Oy	1605	caacacccactacagtctcagacggcaccgtattctgcagagggcagcg99ggggccgcca	1664
Dd	68	CACATCCCACTAATGATGCTGCGACGGCACCGTATCTGCACTGGGCGACGGGGGCGGCCA	127
Oy	1665	gacagcgagactgggaaggatgtaggagcgagactgcagacgaaggaagggaccactatgcg	1724
Dd	128	AACACAGCACTACTGGGAGATGAGAGACGAGCTGCATACGAGGCGGGGAGACCCATMGCG	187
Oy	1725	aggaggaatgscacacaccccccgagctgtgtgtgtggagctagccccttgacacacac	1784
Dd	188	AGGAAGAATGGCCACACCCCAGGCAAGTGTGTGTGTGAGGCACTATCCCCCTTGACACACAC	247
Oy	1785	acacagacacacacactactcgtatgcatgttatgtcacacacatgvcgcgcacacgtctcc	1844
Dd	248	ACACAGACACACACACTACCTGATGATGCAATGTATGCACACACTATGCGCGCACACGTGCTCC	307
Oy	1845	ctgaaggcacacacgtatgcacacacgcacacatgcacagatatgcgccttgacacacagata	1904

Db	308	CTGAAGGCAACAGTACGGACACACGACACATGACAGATATGCGCTGTGGGACACAGATA	367
OY	1905	agctgcaccaatgtcacgcgcacacagacagatgtccagaacatacaagaagatgtctgc	1364
Db	368	AGCTGAGCC---TGGACGCGACACGGACACAGAGACATGCGCACAAATACAAAGGACATGCGCTGC	424
OY	1965	tgaacatcacacgcacaccccatgtcgcaagatgtctgtcctgtgacacacacacacacacg	2024
Db	425	TGAACATACACACGACACCCATCGCAGATGTCTGACTGTGGACACACACACACACACGG	484
OY	2025	atatgtctgtcgtgacgcacacacgtgtcagatatgtatctcgtgacacacacgtgtcacagat	2084
Db	485	ATATGCTTCTCTGGAGCGACACACACTGTCAGATATAGATCCGGACACACACAGTGTGACAGAT	544
OY	2085	atgtctgtcgtgacacacagataatgtctgtcctgtgacacacacatgtcacgatatgtcctg	2144
Db	545	ATGCTGCTGTGACACACAGAT--TGTCTCCTTTGACACACACATGTCACGAGATATGCGCTGC	602
OY	2145	acacacacacacacacacgtgtgtcacagatatgtctgtcgtggaacgcacacacatgtcacata	2204
Db	603	ACACACACACACACACACAGTGC--CAGATATGCTGTGGCTGGACACGC--CACACATGCCAATA	660
OY	2205	tgtctgcctgtggaacacacattccagacacacgtgtgacagagcgacagatatgtctgctgtgaca	2264
Db	661	TGCTGTGCTGGGCACACACTTMCAGACACACCTGTGCACAGCGCAATTTTGTCTGGCGGACCA	720
OY	2265	cacgcgatactatcgtctctagtcacacacacacgcagacatgtctgtccggaacacacacg	2324
Db	721	CACGAGATATGTGTGGGTATGTACACACACACGACAGATGTCTGTGTGGGACACACACG	780
OY	2325	catgtcacagatatgtctgtcgcg	2345
Db	781	CA-GCACAGATATGTGTGGCGG	800
RESULT	7		
LOCUS	AW411525/c		
DEFINITION	fn12g11.y1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2964693 3', mRNA sequence.		
ACCESSION	AW411525		
VERSION	AW411525.1	GI:6937066	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 697)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: rgs@bbs.fremail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC) Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www.bio.lnl.gov/bdrp/image/image.html Plate: LCM64 row: N column: 22 Seq primer: M13RPL reverse primer (ABI). Location/Qualifiers 1..697 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2964693" /clone_lib="NIH_MGC_17" /tissue_type="rhadomyosarcoma" /lab_host="DH10B (phage-resistant)" /note="Organ: muscle; Vector: pORF7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming.		



QY 2289 acacacacgacagatgctgctccgacacacacacgcatgacagatatgctgtccgac 2348  
||  
DB 481 AC-----AC 484  
QY 2349 acacacacgacagatatgctgctccgacacacacacagatatgctgtccacac 2408  
|||||  
DB 485 ACACACACGACCGACAGATATGCTGCTGACACACACACACAGATATGCTGCTACACT 544  
QY 2409 cacacacgctgacagatatgctcctgacacacacatgacagatatgctgtcctgacatg 2468  
|||||  
DB 545 CACACACGTCAGATATGCTGCTGACACACACAGTGCACAGATATGCTGCTGACATG 604  
QY 2469 cacacacgctgacagatatgctgctccgacacacacacacacacacacacacacacac 2577  
|||||  
DB 605 CACACACGTCAGATATGCTGCTGACACACACAGTGCACACAGTGCACAGATATGCTGCTG 664  
QY 2528 ctggac 2586  
|||||  
DB 665 CTGGGACACACTTCGCGACACACATGCACACACAGAGTGCAGATATGCTGCTGACAC 724  
QY 2587 acgacacgctgac-gtgctcttggaagggtgctgcgtgaagcctgacgctgctgctg 2645  
|||||  
DB 725 ACCGACAGTGCAGATGCTGCTGGAAGGTGTGCGGGAAGCTGATGCTGCTGCTG 784  
QY 2646 aggc 2649  
|||  
DB 785 AGGC 788

RESULT 9  
LOCUS BE722083 694 bp mRNA EST 26-OCT-2000  
DEFINITION 601140878f1 NIH\_MGC\_9 Homo sapiens cDNA clone IMAGE:3140874 5',  
mRNA sequence.  
ACCESSION BE722083  
VERSION BE722083  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 694)  
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: DCTD/DRP  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at: Image.lnl.gov  
Plate: LNCM109 row: k column: 19  
High quality sequence stop: 664.  
Location/Qualifiers  
1. 694  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3140874"  
/clone\_1lb="NIH\_MGC\_9"  
/tissue\_type="adenocarcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: ovary; Vector: pOT7; Site:1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCACGCG(c). Size selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the Laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."  
BASE COUNT 206 a 224 c 153 g 111 t  
ORIGIN

Query Match 20.2%; Score 644.2; DB 165; Length 694;  
Best Local Similarity 99.4%; Pred. No. 1.9e-137;  
Matches 657; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
QY 1853 acacgctgac 1912  
|||||  
DB 1 ACACAGTACGAC 60  
QY 1913 aaatgac 1972  
|||||  
DB 61 GAATGACACGAC 120  
QY 1973 cac 2032  
|||||  
DB 121 CACACGAC 180  
QY 2033 tctggaac 2092  
|||||  
DB 181 TCTGACGAC 240  
QY 2093 tgac 2152  
|||||  
DB 241 TGGAC 300  
QY 2153 acacacacgctgac 2212  
|||||  
DB 301 ACACACACGCTGAC 360  
QY 2213 ggaac 2272  
|||||  
DB 361 GGACACACACTTCGCGAC 420  
QY 2273 tatgctgtcgtgac 2332  
|||||  
DB 421 TATGCTGTCTAGTAC 480  
QY 2333 gatatgctgctgac 2392  
|||||  
DB 481 GATATGCTGTCTGCGAC 540  
QY 2393 atgctgctcacaac 2452  
|||||  
DB 541 ATGCTGCTCTCAAC 600  
QY 2453 atgctgctgac 2511  
|||||  
DB 601 ATGCTGCTGTGAC 660  
QY 2512 c 2512  
|  
DB 661 c 661

RESULT 10  
LOCUS BE745264 1224 bp mRNA EST 15-SEP-2000  
DEFINITION 601574033f1 NIH\_MGC\_9 Homo sapiens cDNA clone IMAGE:3834921 5',  
mRNA sequence.  
ACCESSION BE745264  
VERSION BE745264  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 1224)  
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgabbs-r@mail.nih.gov





```

VERSION      BG31236.1  GI:13137674
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 983)
              NIH-MGC http://mgi.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              JOURNAL
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: DCRD/DNP/Gazdar
              CDNA Library Preparation: Ling Hong/Rubin Laboratory
              CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNLN at:
              http://image.llnl.gov
              Plate: L1CM1241 row: 9 column: 14
              High quality sequence stop: 618.
              Location/Qualifiers
                1..983
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:4549669"
                /clone_lib="NIH-MGC_18"
                /tissue_type="large cell carcinoma"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: lung; Vector: pORF7; Site:1: XhoI; Site:2:
                EcoRI; CDNA made by oligo-dT priming. Directionally cloned
                into EcoRI/XhoI sites using the following 5' adaptor:
                GGCAGAG(G). Library constructed by Ling Hong in the
                laboratory of Gerald M. Rubin (University of California,
                Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                Superscript II RT (Life Technologies). Note: this is a
                NIH-MGC library."
BASE COUNT   222 a 305 c 317 g 139 t
ORIGIN
Query Match 19.5%; Score 622; DB 152; Length 983;
Best Local Similarity 95.8%; Pred. No. 2.5e-132;
Matches 736; Conservative 0; Mismatches 20; Indels 12; Gaps 9;
QY 170 ggaacaaagatgagcgcgacatccaca-ggggtgtgagcgcgttcggtgtgcgcgcag 228
    |||||||
DB 2 GGACCAATGATGCGCCGACCATCCACAGCGGCTGAGCCGCTTCGCGCTGCGCGAG 61
QY 229 gggctgaagtgaaagcaggtgagcgaggaagatgacgagcgtgtgacgtgtgcaagccac 288
    |||||||
DB 62 GGGCTGAAGTGAAGCAGGAGTGAGCGGAGAGATCCGGCGGTGATGATGCAAGGCGCAC 121
QY 289 aacggtctgagcgcgtgagcgtcaactacacccgtgctgtgtgacatgaagccca 348
    |||||||
DB 122 AACGGCTTCGGGCGCTGAGCGTCACTACACCTCTGCTGCTGATGACATTAGCCCA 181
QY 349 ggggaagagagagcctggggcccgacagctctctgtgggggtgaagaagaccggcagccaa 408
    |||||||
DB 182 GGGAGAGAGAGCCT-TGGCCCGAGAGCTCTCTGGGGGTAAAGAGACCCCGCAGCCAG 240
QY 409 caatggcagcagaccgctctacacagccctccaagatgaagcgccggtgtgacgacag 468
    |||||||
DB 241 CAGTGGGACAGCAGCGCGCTTCACACAGCCTCCAAATGAGGCGCGGATGATGCGACGG 300
QY 469 ccggtgggtagctcgtggtggtcctaagtgtgtgacgaggggacacccgtgcccgaatc 528
    |||||||
DB 301 CCCGTGGGTAGCTCTCTGCGGCTCAAGTGGGTGGCCAGCGGACACCTCGGCCGACATC 360
QY 529 acgtgagtgaaagagcagcagagccttgacggtcccaagagccgctgagcccaagaag 588
    |||||||
DB 361 ACGTGATGAAGAGCAGACGAGGCTTGAACGCGCCAGAGGCGCGCTGAGCCAGAGAGAG 420

```

```

QY 589 aagtggacactgaagcctgaagaacctgtgcgagcagagacagcggcaaatataacctgtgcgc 648
    |||||||
DB 421 AAGTGAACACTGAGACCTCGAAGAACTCTGGCGCGGAGAGACAGCGG-AAATACACTGCGCG 479
QY 649 gtgtgcaacgcgcggcgccatcaacgcaactacaaggtgtgattgtatccag-cggac 707
    |||||||
DB 480 GTGTGCAACCGCGCGGCGCGCATCAACGCCACCTACAGGTGATGTGATTCAGACGAGAA 539
QY 708 cccgttcaaacgcccgtgtgtcaagcagcagcccccgtgtgtgtgtgtgtgtgtgtgtgtgtgt 767
    |||||||
DB 540 CCGTTTCAAGCCCGTGTCTCACAGCAGCAGCATCCCTGTAACACAGCGGTGGACTTGGGGGG 599
QY 768 gacacgtctcttccagtgaaggtgctgcagcagcagcgtgaagccggtgtgtgtgtgtgtgt 827
    |||||||
DB 600 GACACAGTCTT-CAGTGAAGAGTGTGCGAGGAGAGCTG-AGCCGCTGAT-CAGTGGCTGAA 656
QY 828 ggcgtgtgagtagcggcgagcggcgagcggcgacacatccacatc-----gagtggcgggc 883
    |||||||
DB 657 GAGCGT-GAGTACGCGCGCGGCGGCGCGCACACACTCCACCTCGAGTGTGCGGCGGCGCC 715
QY 884 agaagttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 931
    |||||||
DB 716 CGAAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 931
    |||||||

RESULT 14
BE378421 753 bp mRNA EST 21-JUL-2000
LOCUS BE378421
DEFINITION 60123755F1 NIH-MGC_44 Homo sapiens cdna clone IMAGE:3609013 5',
mRNA sequence.
ACCESSION BE378421
VERSION BE378421.1 GI:9323886
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 753)
              NIH-MGC http://mgi.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              JOURNAL
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: ATCC
              CDNA Library Preparation: Ling Hong/Rubin Laboratory
              CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNLN at:
              http://image.llnl.gov
              Plate: L1CM262 row: m column: 14
              High quality sequence start: 7
              High quality sequence stop: 582.
              Location/Qualifiers
                1..753
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:3609013"
                /clone_lib="NIH-MGC_44"
                /tissue_type="endometrium, adenocarcinoma cell line"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: uterus; Vector: pORF7; Site:1: XhoI; Site:2:
                EcoRI; CDNA made by oligo-dT priming. Directionally
                cloned into EcoRI/XhoI sites using the following 5'
                adaptor: GGCAGAG(G). Library constructed by Ling Hong
                in the laboratory of Gerald M. Rubin (University of
                California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT 172 a 234 c 199 g 148 t
ORIGIN
Query Match 18.8%; Score 598.6; DB 166; Length 753;

```



ACCESSION	BG326224
VERSION	BG326224.1 GI:13132661
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Catenata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 727) <a href="http://mgs.nci.nih.gov/">NIH-MGC http://mgs.nci.nih.gov/</a> . National Institutes of Health, Mammalian Gene Collection (MGC)
AUTHORS	Unpublished (1999)
TITLE	Contact: Robert Strausberg, Ph.D.
JOURNAL	Email: cgaps@email.nih.gov Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Plate: LNCMI275 row: h column: 09 High quality sequence stop: 722.
COMMENT	location/Qualifiers
FEATURES	1..727
SOURCE	/organism="Homo sapiens" /db_xref="taxon:9606" /clone_image:4562744 " /clone_lib="NH_MGC_14" /tissue_type="renal cell adenocarcinoma" /lab_host="DH10B (phage-resistant)" /note="Organ: kidney; Vector: pOTR7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGGCAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)." ."
BASE COUNT	137 a 244 c 233 g 113 t
ORIGIN	
Query Match	18.7%; Score 595.6; DB 152; Length 727;
Best Local Similarity	94.8%; Pred. No. 2.7e-126;
Matches	681; Conservative 0; Mismatches 29; Indels 8; Gaps 6;
DQ	468 gcccgctggtagtcgccgctgcgaagtgcgtgtggccacaggcaacctcggcccgaact 527 
DQ	7 GCCCGGTGTACTCCGTGCGTGCAAGTCGCGGCCAACCGGCCACCCTCCGG-CCTACAT 65 
OY	528 cacgtgatgtaagaagaccagaccccttgacgcgcgcccaagagccgctgaagccaaagaaga 587 
DQ	66 CACGTGATGAAGAAGACGACACAGGCTTTGACGCGCCCCAAGAGCGCGCTGACGCCCAAGAA 125 
OY	588 gaagttggaacttgaagcctctaagaaccttcgcgcccggaggaacgcgccaataaaccttcgcg 647 
DQ	126 GAAGTGAACAATAAGCTGAAGAACTGCGGCGCGAGGAGCACGGCGAATAATACACTCGCCG 185 
OY	648 cgtgtcgaacccgcgcgcgcgcacaacacccacacttaacaagtgatgtatatcacaggagac 707 
DQ	186 CGGTGTGAACCGCGCGGCGCCCATCAACGCACTTAACAGGTGATGTATCCAAGCGGAC 245 
OY	708 ccgtccaagccccgtcctcacaagacagcaacccccggaaacaagaagttggaacttcggggg 767 
DQ	246 CCTTCCAAACCCGCTGCTCAAGGACGACGCCGCCGCGGAACACGACGAGTGGATTCCAGTGG 305 
OY	768 gaccacgtctcttcagttgcaaagtlgcgcagcgaagcgtgaagccggtgatccattgcttgaa 827 
DQ	306 GACCAACGTCTTCCAGTAGCAAGGTGGCAGACGAGAGTGAACCCGCTGATCACTAGTGGTGA 365 
OY	828 gcgcgtgcgagtaaagccgcgcgcgcgcgcaccaactccacaatcatgatgttgtgcgcgaaga 887 
DQ	366 GGCGGTGAGATTACGGCGCAGGGCCGACCAACTCCACCATGATGTGGCGGCCGACGAA 425 

QY	888	gttttgtagctgcgcaagagtacgtgtgttcgcggccgaagcgtcttacttaataa	947
Db	426	gttttgtgtctctgccaccagggtgacgtgtgtgtcgcg--ccgacggctcttacctaataa	484
QY	948	gtctcatatcacccgtgtcccgcagaattgcgagatgatcatctgtcttgcgcga	1007
Db	485	gtctctatcatcccggtccccgcagagactgcccgaatgcccgaatgatctgttggcgcga	544
QY	1008	caccatggtgtacagcttccgcagagcctctctcaacgtgtctgcgaacccaacgcc	1067
Db	545	caccatgggctacagcttccgcagcggccttctctcaccgctgcacaccaaaaaccca	604
QY	1068	aggcgcacacctgtgtgccttcctc-gtccctcggccactagctgcgcgttgcctgtcatcg	1126
Db	605	gttgcgcacctgtgtgccttctcagttctctgcgaccttaccttgccttgc-ccgtgtctatnccg	663
QY	1127	gcatccacgcgcgcgtctcttcatctcttgcgacacctgtctctgtgtgttgcgcagc	1184
Db	664	G-ATTCAGGCGGGGTCTTCAATCCG--ACCCGTCCTCCTGTGCCTGCCAGGCC	717
RESULT	17		
BG332879			
LOCUS	602430724F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4548701 5'	EST	27-FEB-2001
DEFINITION	mRNA sequence.		
ACCESSION	BG332879		
VERSION	BG332879.1 GI:13139317		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheraia; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 735)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs@mail.nih.gov		
	Tissue Procurement: DCRD/DTP/Gazdar		
	CDNA Library Preparation: Ling Hong/Rubin Laboratory		
	CNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.llnl.gov		
	Plate: LNCM1238 row: o column: 06		
	High quality sequence stop: 702.		
FEATURES	Location/Qualifiers		
source	1..735		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4548701"		
	/clone_id="NIH_MGC_18"		
	/tissue_type="large cell carcinoma"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: Lung; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."		
BASE COUNT	223 a 234 C 164 G 114 T		
ORIGIN			
Query Match	18.7%	Score 594.8	DB 152; Length 735;
Best Local Similarity	94.3%;	Pred. No. 4,2e-126;	
Matches .696;	Conservative	0; Mismatches 27;	Indels 15; Gaps 7



QY	1729	ggatattggccaggaacccccaggaagctcgtgtgtgtgtaggacatagcccccttggaaacacacacac	1788
Db	1	GGAAATGGCCAGGACACCCCAAGCAATCTGTGTGTGAGGCAATAGCCCTGGAGACACACACAC	60
QY	1789	agacacacacactactctgtgtatgtcatlalyaacacacatgcgcgcacacagtgctccctga	1848
Db	61	ACACACACACACTACTCTGGATGTGATGTATGACACACACATGCGGACACACGTGCTCCTGA	120
QY	1849	aggcacaagctacgcacacacacgcacatgycacagaatalgcgcctctgggcaacagaataagct	1908
Db	121	AGCGACACGTAGCGACACACAGCAGCATGACAGATGATGCGCGCTGGGACACAGATTAAGCT	180
QY	1909	gcccaaatgcacgcgcacacgcgcacagagacatgycgcagaacataacaaagacatgctgtctgaa	1968
Db	181	GCCCAAAATTCACAGCCACACCCACACAGACATGCTCCAGAAACATACAAAGACATGCTGCCGTA	240
QY	1969	catacacagcacaccccacatgcgcgcagatgtgtctccctggacacacacacacacagatat	2028
Db	241	CATPACACACGACACCCCAATGCGAGATGTGCTCTGGAGACACACACACACACAGGATAT	300
QY	2029	gctgtctggacgcacacaca-cgtgcagataatgtatcgcgcgcgcacacacacgltgcacagatatg	2087
Db	301	GCTGTCTGGACGACACACACCGTGCAGATATGTGATTCGCGACACACACGTGCACAGATATG	360
QY	2088	ctgcctgcgcacacacagataatgctgcctctgcacacacacatgycagatatgtctctgaca	2147
Db	361	CTGCCTGGACACACACAGATATATGCTGCTTGACACACACATGCACGGATATTTGCCCTGGACA	420
QY	2148	cacacacacacacgcgtgtgcacagatatgtgtctgttcgcgcacacacacacatgcagatatgc	2207
Db	421	CACACACACACACGCGCTGCACAGATATGCTGTGTGACACAGCACACACACATGCAGATATGC	480
QY	2208	tgcctgcgcacacacactccacagacacacgltgcacagcgcacg-atalgctcctgcgcacaca	2266
Db	481	TGCGTGGACACACACACTTCCAGACACACGCTGCACAGGCGGAGATATGTGCTGCGGACACA	540
QY	2267	cgcgcgatatgctgtt-ctagtcacacacacacacgcgcagacatgctgtccgcgcacacaca--	2322
Db	541	CCGACGATATGCTGTACTACTATCACACACACACGACGACATGCTGCTCGGACACACACACG	600
QY	2323	-----cgactgcacagatatgctcgtctccgcgcacacacacgcgcgcgcagacatgctgcctcg	2377
Db	601	CAATGTCACCCAGAAATATCTGTTCTCGGAGACACACACACGACGACGAT-TGCTGTCTGG	659
QY	2378	acacacacacagataatgctgcctcccaacactcacacacgtgcagatatgtcctggacaca	2437
Db	660	ACACACACACACATATGCTGTC--TCAAATTCACACACGTCGACGATATGCGGTGGACACA	717
QY	2438	cac-atgtgcacagatat 2454	
Db	718	CACAAATGTGCACAGATAT 735	
RESULT 18			
LOCUS	BG423479	957 bp	EST 14-MAR-2001
DEFINITION	6024493337F1 NIH_MGC_14	Homo sapiens cDNA clone IMAGE:4587752	5'
ACCESSION	BG423479		
VERSION	BG423479.1	GI:13329985	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostom		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a>		
	Tissue Procurement: DCTD/DTF		
	cDNA Library Preparation: Ling Hong/Rubin Laboratory		





Db	582	TATCTGTCTAGTACACACACGACAGCAACTGCTGTCCGGACACACACAAAGCATGCCA	641
Qy	2331	caga--tatgtcgtccggacacacacagcagcgc	2362
Db	642	CCAAGATTATGCTGTCCGGACCACACACACACGCAC	676
RESULT	22		
Bf340012			
LOCUS			
DEFINITION	Bf340012	974 bp mRNA	EST 22-NOV-2000
ACCESSION	60203694Bf1 NCI_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4185059		
VERSION	5', mRNA sequence.		
KEYWORDS	Bf340012		
SOURCE	Bf340012.1 GI:11286474		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	1 (bases 1 to 974)		
COMMENT	NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov Tissue Procurement: David N. Louis, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMN at: http://image.llnl.gov Plate: LHAM9503 row: o column: 12 High quality sequence stop: 603. Location/Qualifiers 1..974 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4185059" /clone_1id="NCI CGAP_Brn64" /tissue_type="glioblastoma with EGFR amplification" /lab_host="DH10B (TI phage-resistant)" /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.57 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP library."		
FEATURES			
source			
BASE COUNT	286 a 277 c 226 g 185 t		
ORIGIN			
Query Match	17.7%; Score 564.6; DB 147; Length 974;		
Best Local Similarity	84.3%; Pred. No. 3.7e-119;		
Matches	685; Conservative 0; Mismatches 119; Indels 9; Gaps		
Qy	2152	cacacacacgcttgacacagatatgctgttgtgaacagcacacatcgagatatgttcgc	2211
Db	10	CGACACACGTGTGCACAATATGCTGTGTGGACAGCACAACAATGCAATATGCTGCC	69
Qy	2212	tggacacacacttcacgaacacgltgcacaaggcgacagatatgctgcttggacacacgcg	2271
Db	70	TGGACACACACTTCGAGACACACAGCTCACAGGCGCAGATATGCTGCTGGACACACGACG	129
Qy	2272	atatgctgtctagtcaacacacacacgacagatagctgctcgcgacacacacacgcatgcc	2331
Db	130	ATAATGCTGTCTAATGCACACACACACGCAACATGCTGCCGACACACACGCGATGCAC	189
Qy	2332	agatatgctgtccggacacacacacgacacagacatatgctcctctggacacacacagat	2391
Db	190	AGATATGCTGTCTCGGACACACACACGCGACGCAATATGCTGCTGGACACACACAGAT	249
Qy	2392	aatgtcgtcctaaacctacacacacgltgcagatatgtccttggacacacacatgtgcacag	2451

Db	Accession	Gene	Chromosome	Start	End	Strand	Length	Source
Db	250	AATGCTGCTCCACACACTCACACAGCGCAGATATTTCCTGGACACACACATGTCGACAGA	309					
Qy	2452	tatgctgctcgtgacatgacacacacgctgcagatattgctctccggatatacaac-cgcaagcgc	2510					
Db	310	TATGCTGTCTGGACATGACACACAGCTTACGATATTGCTGTCCTGGATATCACACCGCAGCGAC	369					
Qy	2511	acatgcagatatgctgctcgtgacacacactccggacacacacatgcacacacagtgacaa	2570					
Db	370	ACATGCGAGATATGCTGCTGGGACACACACTTCGGGACACATGCAACATGACACAGTGCAGA	429					
Qy	2571	tatgctgcctgcgaacacacagacactgcagctgctctttggagaggctgctgcgtgaagcctgc	2630					
Db	430	TATGCTGCCTGCAGACACACGACAGATGACGTGCTTTGGAGGGGTGTGCCGTGAAGCCTGC	489					
Qy	2631	agtaacgtgctgcgtgaaagctcactagttatgatagaggactctcctgtcactccgtaactcc	2690					
Db	490	AGTACGTGTGCGCGTAGGAGCTCATAGTTGATGAGAGGAGACTTTCCTGCTCCACCCCTCACTCC	549					
Qy	2691	cccaactctgcgcgcctctgtcccccgtcctaagctcccccgtcctccatcccccgtcctcc	2750					
Db	550	CCCAACTCTGCGCGCTCTGTCTGCTCCGCG-TCGTCCCGCGTCCATCCCGCTCTGTCC	604					
Qy	2751	ctggagcttgagcgctatttttgccacactgctcttggtgcccagaagttccct-actgctg	2809					
Db	605	CTGGCTTGCGGTATTTTTGCAAGCTGAC-ATGAGGTTGCCAGTGAAGTCCCTTAACGCTG	662					
Qy	2810	tgaggctggggttgagggcacacagcagccccaagccttgagaggtcgtgagcccatggtagt	2869					
Db	663	GGGGATGGGATGGCGGCAAGAACGACCACCAAGCAATTGGAAAGGTGGAGCCATGGCAAGG	722					
Qy	2870	gctacatcccaactgcattctcccccctgacacagagagaaggcgcttgattatattaa	2929					
Db	723	CTAATCCCAATGCATATTCCTCCGAGAACAAAGAAATGGCCTTGGAATTATATTAAGATGAA	782					
Qy	2930	gaatgaagaatataatataatataatgaagga	2962					
Db	783	AATATATATACACGATGGAAGATATATATGCGCGTGA	815					
RESULT	23							
LOCUS	BC422503							
DEFINITION	60244507BP1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4587786 5',							
ACCESSION	BC422503							
VERSION	BC422503.1							
KEYWORDS	EST.							
SOURCE	human.							
ORGANISM	Homo sapiens							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.							
TITLE	1 (bases 1 to 945)							
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.							
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)							
	Unpublished (1999)							
	Contact: Robert Strausberg, Ph.D.							
	Email: cga@bbs.fda.gov							
	Tissue Procurement: DCTD/DTF							
	cDNA Library Preparation: Ling Hong/Rubin Laboratory							
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)							
	DNA Sequencing by: Incyte Genomics, Inc.							
	Clone distribution: MGC clone distribution information can be							
	found through the I.M.A.G.E. Consortium/LNL at:							
	http://image.lnl.gov							
	Plate: LICM1318 row: k column: 19							
	High quality sequence stop: 669.							
FEATURES								
source	location/qualifiers							
	1..945							
	/organism="Homo sapiens"							
	/db_xref="taxon:9606"							
	/clone="IMAGE:4587786"							
	/clone_id="NIH_MGC_14"							
	/tissue_type="renal cell adenocarcinoma"							

/lab\_host="DH10B (phage-resistant)"  
/note="Organ: Kidney; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adapter: GCCACGAG(5'). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."  
BASE COUNT 198 a 295 c 334 g 118 t  
ORIGIN

Query Match 17.5% Score 558: DB 153: Length 945:  
Best Local Similarity 91.1% Pred. No. 1.2e-117:  
Matches 696; Conservative 0; Mismatches 50; Indels 18; Gaps 9;  
QY 215 ggcgtctgcgcgagggcgtgaaggtgaagcaggtgaagcgaggaagatgcgcgcgtgaag 274  
DB 2 GCCTGCTGCCGCGACTGCTTAAGGTGAAGCAGGTGAGCGGAGATGCCGCGCTGTACG 61  
QY 275 tctgcagagccacacacgcgtctgcgcgcgtgaagccttaacacacccctgcgtctg 334  
DB 62 TGTGCAAGCGCACCAACGCGCTCGGCGAGCTGAGCGTCACTACACCCCTCGTGTGCTGG 121  
QY 335 atgaacatagcccaaggaagagaagcctggtgcccgaagcctcctctggtggtcaagag 394  
DB 122 ATGACATTACCCA-GGAAGAGAGAGCTGGGGCCGACACTCTCTGTGGGGTCAAGAG 180  
QY 395 accccacacacacacagcttgagcagcagcgtctcaacacacccctcaagaatgaagcgc 454  
DB 181 ACCCGCGCACCGACAGTGGGCGACGCGGCTTCACACAGCCCTCCAGATGAGGCGCC 240  
QY 455 gggctgacgcagcgcgcgttggtgagtcgcgtgcgcctcaagctgcgtgcgcgcgcgc 514  
DB 241 GGGTGATCCGACGCGCGCGGTGGTGGTCCGTGGCGGTCAAGTGCGTGGCGAGCGGCGAC 300  
QY 515 ctgcgcgcgcacatcagctggaaggaagcagcagccttgagcgcgcgcgcgcgcgcgtg 574  
DB 301 -TCGGCCGACATCACCTGTGATGAGGACACGAGCGCTTGACGCGCGCCAGAGCGCG -TG 357  
QY 575 agcccgaggaaggaagtgagcactgaagcctgaagacgcgcgcgcgcgcgcgcgcgcgc 634  
DB 358 AGCCCGAGGAAGAAAGTGGACACTGAGCTTGAAGAACTGCGGCGCGGAGGACGCGGCA 417  
QY 635 aatacactgcgcgcgttgctgaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 693  
DB 418 AATAACAACCTCGCGGTGTCGACCGCGCGCGCCATCAAGCCGCACTCAAGGTGAT 477  
QY 694 gtagatcagc 753  
DB 478 GTATATCCAGGGGACCCGTTCAAGCCGTTCTCAAGGCAAGCAGCCCGTGAACACACG 537  
QY 754 gtagacttcgggggagaccagctccttcacagtgaaggtgagcagcgcgtgaagcgcgtg 813  
DB 538 GTGGACCTTGGGGGGGACAGGCTCTTCACATGCGAAAGTGCAGAGGAGTGAAG-CGGTG 596  
QY 814 atccagctgcgtgaagc-gcgttgagtagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 872  
DB 597 ATCCAGTGGCTGAAGCGCGCTGAGTACGCGCGCTTACG---CCGCAAACTCCAGACATCGA 653  
QY 873 tctggtgc 930  
DB 654 TGTGGGCGCGCAGAAAGTTGTGGGTGCTTCCACAGGCGTACAGGTGTGGCAGCGGC--- 709  
QY 931 ggtcctcactaataagctgcctcatcaccgtgcccgcgcgcgcgcgcgcgcgcgcgcgc 974  
DB 710 --GAGGTCAACTAAAGGTGTCAACACCGTCCGCGCAGAGA 751  
RESULT 24  
BE873767  
LOCUS BE873767 913 bp mRNA EST 20-OCT-2000

DEFINITION 601483916F1 NIH\_MGC\_69 Homo sapiens cDNA clone IMAGE:3886683 5',  
mRNA sequence.  
ACCESSION BE873767  
VERSION BE873767.1 GI:10322543  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 913)  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)  
UNPUBLISHED (1999)  
CONTACT: Robert Strausberg, Ph.D.  
EMAIL: cgabs-remail.nih.gov  
Tissue Procurement: DCTD/DTF/Gazdar  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
plate: LLAM9663 row: 0 column: 04  
High quality sequence stop: 648.  
FEATURES  
source  
1..913  
location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3886683"  
/clone\_lib="NIH MGC 69"  
/tissue\_type="large cell carcinoma, undifferentiated"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: Lung; Vector: PCMV-SpR6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.1 kb. Library constructed by Life  
Technologies."  
BASE COUNT 216 a 259 c 253 g 185 t  
ORIGIN  
Query Match 17.3% Score 552: DB 141: Length 913:  
Best Local Similarity 95.6% Pred. No. 2.9e-116:  
Matches 623; Conservative 0; Mismatches 20; Indels 9; Gaps 5;  
QY 2415 cgtgcagatatctgcttgacacacacatgacacagatgctgctgcgcgcgcgcgcgcgc 2474  
DB 1 CGTGCAATATTTGGCTGGACACACAGTGGACAGATGCTGTGGACATGACACACA 60  
QY 2475 cgtgcagatatgctgctgc 2534  
DB 61 CGTGCAATATTTGGCTGGACACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 120  
QY 2535 cacacttcgc 2594  
DB 121 CACACTTCCGGACACATGACACACAGTGCAGATATGCTGCTGGACACAGCAGAC 180  
QY 2595 tgaacgtcttctgggaaggtgctgcgtgaagcctgcgcgcgcgcgcgcgcgcgcgcgcgc 2654  
DB 181 TGACGTGC-TTTGGAGAGGTGTCGCGTGAAGCGTGAAGCGTGAAGCGTGAAGCGTGA 239  
QY 2655 gttgaggaaggaacttccctgc 2714  
DB 240 GTTGATGAGGAGACTTTCCTGCTCCACAGTCCACACCTCCCAACTCTGCCGCTGTGCC 299  
QY 2715 cgcctcagctccgc 2774  
DB 300 CGCCTTCACTCCG-CTTCACTCCCGCCTCTGTCTCTCTGCGC--TTGCGGTATTTTGGC 356  
QY 2775 acctgccttggtgc 2834  
DB 357 ACCTGCTTGGGTGCCAGAGAGTCCCTACTGCTGTGGGTGGGGGTGGGGGCGACACAG 416  
QY 2835 ccccaagcctgagagcgtgagcccatgctagtgctcatcccaatgcatctccccc 2894

|||||  
Db 417 CCCCAGGCTGAGAGCTGGAGCCCATGGCTAGTGGCTCATCCACAGCTTCCTCCCC 476  
Oy 2895 tgacacagagaaggagccttgatattatattagaataagaataataataatga 2954  
Db 477 TGACACAGAGAGGAGGCTTGGTATTATTATTAAGAAATGAATATATAATGA 536  
Oy 2955 tggaaagaagactgggttcgaaggagctgtctctcctggggccggagaccgctgt 3014  
Db 537 TGGAGGAAGAGACTGGGTGGAGGAGCTGTCTCTCTCCCTGGGGCCGGAGCCGC 592  
Oy 3015 ctttaagcattgctatgacacacacccctgcaggccagacacacacccca 3066  
Db 593 CTTTCAG-CATGCTGATGACACATCCGTTTCAGGAGAAACACCCACCA 643

RESULT 25  
BF970020 784 bp mRNA EST 22-JAN-2001  
LOCUS 602272864F1 NIH\_MGC\_84 Homo sapiens cDNA clone IMAGE:4360958 5'  
DEFINITION mRNA sequence.  
ACCESSION BF970020  
VERSION BF970020.1 GI:12337235  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 784)  
NIH-MGC http://mhc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL0002 row: p column: 15  
High quality sequence start: 3  
High quality sequence stop: 682.  
Location/Qualifiers  
1..784  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4360958"  
/clone\_1lb="NIH\_MGC\_84"  
/tissue\_type="adrenal cortex carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site: 1;  
Nct1; Site\_2: SalI; Cloned unidirectionally; oligo-dr  
primed. Average insert size 1.229 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

BASE COUNT 134 a 267 c 256 g 127 t  
ORIGIN

Query Match 17.0%; Score 543.2; DB 172; Length 784;  
Best Local Similarity 95.2%; Pred. No. 3e-114;  
Matches 636; Conservative 0; Mismatches 23; Indels 9; Gaps 7;

Oy 433 cagccctcaagatgagcgcggtgatcgacggccc-gtggtagctccgtgggct 491  
Db 2 CAGCCCTCCCAAGATGAGCGCGGTGATGCGACGGACCGGTGAGCTCCGTGGGCT 61  
Oy 492 caagtgcgtgacgagcgacccctggccgacatcaagtgtgaagaagcagcag 551  
Db 62 CAAGTGGCTGCGGACGCGGCACTCTCGG-CCGACATCAGCTGATGAAGACGACGAGC 120

Oy 552 ctfgacgcccccaaggccgctgagccccaggaagaatggacactgagcctgaaga 611  
Db 121 CTTGACGCGCCAGAGAGCCGCTGAGCCAGAGAAAGAAATGAGACATGAGCCCTGAAGAA 180  
Oy 612 cctgagcgaggaagacagcggaataaacactgcgcgtgtcgaaccgcgcgccat 671  
Db 181 CCTCGCGCGGAGACAGCGGCAATATACCTGCGCGGTTCGAACCGCGCGGCGCAT 240  
Oy 672 caagcgaactacaagtgtgatgtgatccgagcagccgttccaaagccgtgtcacag 731  
Db 241 CAAGCGCACTTACAAGGTGATGTGATTCAGCGAGGACCGGTTCAGGCCCTGCTCACAGG 300  
Oy 732 caagcaccgcctgacacagacgagtgagacttcggggaggacacgctccttcagtgaaagt 791  
Db 301 CACGACCCCGCTGAACACGAGCGTGGACTTGGGGGACACGCTTCCAGTCAAGGT 360  
Oy 792 gcgcagcagctgaagccggtgatccagtgatcagtgacgctgagtgagtgagcgcgaggg 851  
Db 361 GCGCAGCGAGCTGAAGCGCGGTGATCCAGTGGCTGAACGCGCTGAGAGTACGCGCGAGGG 420  
Oy 852 ccgcacacacccacacacatgattgtggcgcgccaggaagtttgtgtgtccacaggtga 911  
Db 421 GCGCGCAAA-TCCACATCGATGTGGCGGCCACAGATTGTGTGCTGCCACAGGGTGA 479  
Oy 912 cgtgtgtcgcgcgcgcagcagcgtctcctacataaagctgtcatacccggt--gccgcg 969  
Db 480 CGTGTGTGCGCGCGCGCGAGCG-TCTTACTTCATTAAGCTGCTCATACCCCTTGCCTCCCGC 538  
Oy 970 caggaagatgagggatgataatctgctgtggcgcaacacacattgggttaagcttcgc 1029  
Db 539 CAGGACGATGGGGGATGTTATCTTCCTTGGGCGCAAGCGCATG6GCTAC-GCTTCCGC 597  
Oy 1030 aggccttcctcaaccgtgtctgcagaccacacacgcagcagcagcactgtgacctcctgc 1089  
Db 598 ACGCGCTTCCTACCGCTGCTGCTG--AGACCAACACCGGAGGCGCACCTGTGGCTCTCGC 655  
Oy 1090 tccctgcgc 1097  
Db 656 TCCCTCGGC 663

RESULT 26  
BF984116 936 bp mRNA EST 23-JAN-2001  
LOCUS 602306805F1 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:4398048 5'  
DEFINITION mRNA sequence.  
ACCESSION BF984116  
VERSION BF984116.1 GI:12386928  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 936)  
NIH-MGC http://mhc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL0099 row: j column: 01  
High quality sequence stop: 715.  
Location/Qualifiers  
1..936  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4398048"



Db	49	CCTATATTATTTATTTTGTTAACATGAAGAAGCATCCTTGCCCTCC	1
RESULT	28		
LOCUS	AW845240		
DEFINITION	OVO-CT0018-011199-042-a06 CT0018 Homo sapiens CDNA, mRNA sequence.		
ACCESSION	AW845240		
VERSION	AW845240.1	GI:7940757	
KEYWORDS	EST.		
ORGANISM	human.		
SOURCE	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 660)		
JOURNAL	Dias Neto, E., Garcia Correa, R., Verjovsky-Almeida, S., Briones, M.R.,		
MEDLINE	Nagal, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,		
COMMENT	Gudman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,		
	Brunstein, J.A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare		
	, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and		
	Simpson, A.J.		
	Shotgun sequencing of the human transcriptome with ORF expressed		
	sequence tags		
	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)		
	Contact: Simpson A.J.G.		
	Laboratory of Cancer Genetics		
	Ludwig Institute for Cancer Research		
	Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,		
	Brazil		
	Tel.: +55-11-2704922		
	Fax: +55-11-2707001		
	Email: asimpson@ludwig.org.br		
	This sequence was derived from the FAPESP/LICR Human Cancer Genome		
	Project. This entry can be seen in the following URL		
	(http://www.ludwig.org.br/scripts/gethtml2.pl?rl=et2-qv0-CT0018-011199-042-a06&t3=1999-11-01&t4=1)		
	Seq primer: puc 18 forward		
	High quality sequence start: 3		
	High quality sequence stop: 659.		
FEATURES	Location/Qualifiers		
Source	1..660		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_lib="CT0018"		
	/dev_stage="Adult"		
	/note="Organ: colon; Vector: puc18; Site_1: Smat; Site_2:		
	Smat; A mini-library was made by cloning products derived		
	from ORESTES PCR (U.S. Letters Patent application No. 199		
	,716 - Ludwig Institute for Cancer Research) profiles		
	into the pUC 18 vector. Reverse transcription of tissue		
	mRNA and cDNA amplification were performed under low		
	stringency conditions."		
BASE COUNT	114 a 253 c 187 g 106 t		
ORIGIN			
Query Match	16.4%; Score 521.6; DB 121; Length 660;		
Best local Similarity	99.1%; Pred. No. 2.7e-109;		
Matches 555; Conservative	0; Mismatches 4; Indels 1; Gaps 1		
Dy	892	gttgagctgcgccacgagtacgttgttcggcgccgaaggcctctactcaataagctg	991
Dy	3	gttggtctgctgcacggcgtgcgtgtgtgtgcgcg-ccgacggccttaccatattagctg	61
Dy	952	ctcataccggtcccgccagacgagtcgaggatctacctgcttgtagccacaacc	1011
Dy	62	cttatatccccgtcccgcacgagatgcgggatgtactttgcttggcccaaacac	121
Dy	1012	atggagtaagcttcgcgaagccttctcctaaccgtgtcgcgaagccaagaacgcaggg	1071
Dy	122	atggcgttacagcttccgacacggccttccctacagctgtgtgcacaacccaacacgcaggg	181

Query Match	16.3%	Score 519.4	DB 147	Length 873
Best Local Similarity	89.3%	Pred No. 8.9e-109		





LOCUS	BC254594	774 bp	mRNA	EST	13-FEB-2001
DEFINITION	602368534.F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4476609 5', mRNA sequence.				
ACCESSION	BC254594				
VERSION	BC254594.1 GI:12764410				
SOURCE	EST.				
ORGANISM	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 774)				
TITLE	NIH-MGC http://mgs.nci.nih.gov/.				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
COMMENT	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D.				
COMMENT	Email: cgapbs-remail.nih.gov				
COMMENT	Tissue Procurement: DCTD/DRP				
COMMENT	cDNA Library Preparation: Life Technologies, Inc.				
COMMENT	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)				
COMMENT	DNA Sequencing by: Incyte Genomics, Inc.				
COMMENT	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:				
COMMENT	http://image.llnl.gov				
COMMENT	Plate: L1AM10304 Row: C Column: 10				
COMMENT	High quality sequence stop: 656.				
FEATURES	Location/Qualifiers				
SOURCE	1..774				
SOURCE	/organism="Homo sapiens"				
SOURCE	/db_xref="taxon:9606"				
SOURCE	/clone="IMAGE:4476609"				
SOURCE	/clone_11b="NIH_MGC_91"				
SOURCE	/tissue_type="adenocarcinoma, cell line"				
SOURCE	/lab_host="DH10B (phage-resistant)"				
SOURCE	/note="Organ: prostate; Vector: pCMV-SPORT6; Site.1: NotI, Site.2: SalI; Cloned unidirectionally; oligo-dr primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."				
BASE COUNT	231 a 242 c 168 g 133 t				
ORIGIN					
Query Match	16.0%; Score 510.2; DB 175; Length 774;				
Best Local Similarity	92.1%; Pred. No. 11e-106; Indels 15; Gaps 77				
Matches 618; Conservative	0; Mismatches 38; Indels 15; Gaps 77				
Oy	1935	acatgcagaacaataaagagacatgctgcctgaacataacacacgcacacccatgacgaga	1994		
Db	3	ACGCGTCGGAACTACAGAGACATGCTGCTGAACATACACACGACACCATGGCCACA	62		
Oy	1995	ctgtctgctggaacacacacacacacaggaatgctgtctggaagcacaacgctgcaga	2054		
Db	63	TGTGCTGCTGGACACACACACACACACAGATATGCTGTCTGGACCCACACAGTGCACA	122		
Oy	2055	tatgtrtatccggagacacacgctgcacagatatgctgcctgcggacacacagataatgctgc	2114		
Db	123	TATGTATTCGGACACACACAGTGCATATGCTGCTGGACACACATATGCTGCC	182		
Oy	2115	ctgaacacacacatgcacggaatatgctctg--acacacacacacacgctgtgcagat	2172		
Db	183	TTTGACACACACATGCAAGGATATTGCTGTGGACACACACACACACAGCGGTGCACAGAT	242		
Oy	2173	atgctgtctggaacacacacacatgcatatgctgtgctgcggacacacacttcacagac	2232		
Db	243	ATGCTGTCTGGACACGACACACATGCAATATGCTGCTGGACACACATTCACAGAC	302		
Oy	2233	acgtgcacagcgacagatatgctgcctgcggacacacgcgatatgctgtctagtcacacac	2292		
Db	303	ACGTGCACACAGCCAGATATGCTGCTGGACACACGAGATATGCTGTATGTCACACAC	362		
Oy	2293	acacgagacatgctgtcggaacacacacacgcatgacagatatgctgtctcggaacac	2352		
Db	363	ACACGAGACATCTGTCTCCGACACACACACGATGACAGATATGCTGTCCGACACAC	422		

OY	2353	acagcgcacgagatatactgtcgttgaaacacacacag-ataatgtctgcccaac-actca	2410
Db	423	ACAGCGACGGCAGAATATCTGCTGTGGACACACACACAGTAAATATGCTCCCTCAACAACCTCA	482
OY	2411	cacagctgcagaatatcttccttggacaacacacatggaacagatatgtctctgaacatga	2470
Db	483	CACACGTGCACAGATTTCGCTTGACACACACATGTGCACAGATAATGCTGTGGACATGCA	542
OY	2471	cacac--gtgcagatatgtctgtccgatlaa-----cacgcacgcacacatatgcagatatg	2523
Db	543	CACACAGTGCAGTAATATGCTGTCCGGATACAGCACGCCGACCATGCACATGCAAGAAATATG	602
OY	2524	ctgtccttggaccacaaa---cttcgcggaacacac-atgcacacacaggatgcagatgcgcgc	2579
Db	603	CTAGCTCTGGGACACACACACTTCCGGACACACAGATGACACACAGTGCAGATATGCTGGC	662
OY	2580	tggacacacagc 2590	
Db	663	CTGGAACACAC 673	
RESULT	32		
LOCUS	BF968611	762 bp mRNA EST	22-JAN-2001
DEFINITION	602271034F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4359105 5'		
ACCESSION	BF968611	mRNA sequence.	
VERSION	BF968611.1	GI:12335826	
SOURCE	EST.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 762)		
JOURNAL	NIH-MGC http://mhc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LML at: http://image.llnl.gov Plate: LLM9998 row: C column: 10 High quality sequence stop: 609. Location/Qualifiers 1..762 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_image="4359105" /clone_lib="NIH_MGC_84" /tissue_type="adrenal cortex carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site: 1; Notti, Site 2: Salt; Cloned unidirectionally; oligo-dT primed. Average insert size 1.229 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC 94 t		
BASE COUNT	160 A 245 C 263 G 94 T		
ORIGIN			
Query Match	15.7%:	Score 501.4:	DB 171: Length 762:
Best Local Similarity	95.7%:	Pred. No. 1.2e-104:	
Matches 558; Conservative	0; Mismatches 21; Indels 4; Gaps 4:		
OY	89	tctcgtagcagagtggccgcgcgttgagccgacacatgacgctgcatgcccagtggaaggggg	148
Db	118	TCCACAGCGAGATGGCCCGGCTGGCGGACACACTGTGGGGCTGCAGTGTCCAGTGAAGGGG	177





QY	1848	aaggacacagctaaagcaacacacgcagacatgtcacagatattgtccgctctggcaacacagataac	1907
Db	240	AAGGCACACGCTAGCGACACACACGACATGACAGATATGCCGCTTGGGACACAGATTAAGC	299
QY	1908	tgcccaaatgc -acgcacacgcacagagacatgtcccaacatcacaaaggaacatgtctgcctg	1966
Db	300	TGCCCAATGCGACGACGACGACGACAGATGTCCGAGACATACAAGGACATGCTGCTG	359
QY	1967	aac -atcacacgcacacacccaatgcagatgtgtctgtccctgcacacacacacacacacgcga	2025
Db	360	AACATATACACACGACACCCCATGCGAGATGTGCTCCCTGGACACACACACACACACGGA	419
QY	2026	tatgtcttcgcgaagcgcacac -cgtgcagatatgttatctcggacacacagctgcacagat	2084
Db	420	TATGCTGTCTGGAGCGACACACGCTGCAGATATGATATCGGACACACACCTGCACAGAT	479
QY	2085	atgtctgcctgcgaacacagataa -tgtcgccttgcacacacacatgcagatattgtcctg	2143
Db	480	ATGCTGCTGTGGACACACAGATTAAGTCTGCTCTGTAACCATATTCAGAGGATATTCCTG	539
QY	2144	gacacacacacacacac -ggtgtgcagagtatgtctgtgcgaacgcacacacatgcgaa	2202
Db	540	GACAAAGACACACACTGTGTGTCACACAGATATGCTGTGGACACGACACACATGACGA	599
QY	2203	tatgtcttcgcgaacacacacatctcagacacacagctgcacagcgacagatatgc -tgtcctg	2261
Db	600	TATGCTGTGCTGGACAAAGACATTCACGAAACAGTGCACAGGCGCAGATATCTGCTGCG	659
QY	2262	acacacgcgcgata 2274	
Db	660	ACACACGCAGAGTA 672	

RESULT	36				
LOCUS	AV0707425	537 bp	mRNA	EST	09-OCT-2000
DEFINITION	AV0707425	ADB Homo sapiens cDNA clone ADBBjF09.5',	RNA sequence.		
ACCESSION	AV0707425				
VERSION	AV0707425.1	GI:10724690			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homalidae; Homo.				
AUTHORS	1 (bases 1 to 537)				
TITLE	Peng,Y., Song,H., Huang,Q., Huang,C., Gu,Y., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N., Qian,B., Liu,F., Qu,D., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Lu,G., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R., Chen,J., Chen,Z. and Han,Z.				
JOURNAL	Homo sapiens cDNA ADB clones				
COMMENT	Unpublished (2000)				
	Contact: Zeguang Han				
	Chinese National Human Genome Center at Shanghai				
	351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai				
	201203, P. R. China				
	Tel: 86-21-50801919(ex.45)				
	Fax: 86-21-50801922				
	Email: hanzg@hgc.sh.cn				
	This clone is available at CHGC in Shanghai.				
FEATURES	Location/Qualifiers				
source	1..537				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="ADBBjF09"				
	/clone_lib="ADB"				
	/tissue_type="Adrenal gland"				
	/dev_stage="Adult"				
	/lab_host="SOLR"				
	/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"				
BASE COUNT	148 a 151 c 119 g 119 t				
ORIGIN					

Query Match	15.3%;	Score 489;	DB 32;	Length 537;
Best Local Similarity	100.0%;	Pred. No. 7.9e-102;		
Matches 489;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Oy	2658	ctgcgcgcctctctgccccgcctaafcccccgcctcatcccccgcctctgtcccccgcct	2755
Db	1	CTGCCCCGCTCTGTGCCCGCTCAAGTCCCGCTCCATCCCGCGCTGTGCCCGCGCT	60
Oy	2758	tgcgcgcataatttgcaccactgccttgcgtccacagagtcacctactgcctgtgcgcctgc	2817
Db	61	TGGCGGCGATTATTGGCCACTGGCTTGAGGCCACGAGAGCCCTACTGCTGGGCTGG	120
Oy	2818	ggttgcggggcaacagcaacccccaaagccttgaaagcttgaaacccatgctagtggtccatcc	2877
Db	121	GGTtGGGGGCAcAcAcACCCCAAGcCCTGAAAGGCTGAGACCCATGGCTAGTGTGGCTATCC	180
Oy	2878	ccactgcattctccccctgcacagaaagagggccttgcattatattaaagaaatga	2937
Db	181	CCACtGCATtTCtCCtCCtTGACACAGAGAAAGGGcCTTGtGATtTATtTAAAGAAATGA	240
Oy	2938	gataataatataatgttggaagaagagactcgggtctgcaggagacttggtctctcctgcgg	2997
Db	241	GATATATtTATATATGATGGAAGAGAAcAGCGGGtTGCAGAGAGCTGGtGCTTCCTCGGGG	300
Oy	2998	ccgcgagaccgcctgcgtccttcacagctgcgtga tgacacacaccccgttcacagcagaaac	3057
Db	301	CCCGGAGCCGCGcCTTGtCTTCAAGcCATGCTGATGACACACCCCGTCCAGGGCAGACAC	360
Oy	3058	cacccccaaccacactgcgtgcgtgcgcacagatctctgtaattttagtagagttgcag	3117
Db	361	CACCCCCACACCCACTGCTCGTGGGGCCCAcGATCTCTGATtTTTAACTAGAGTTTGAAG	420
Oy	3118	ctgaagcccccgtatatattatattctgttgaacatgaagctgaaccccttccctcaac	3177
Db	421	CTGAGAGCCCGTATATTTATTTATTTTGTtAAACATGAAGAGCATCCTTTCCCTCCAA	480
Oy	3178	aaaaaaaaa 3186	
Db	481	AAAAAAAAA 489	

```

RESULT 37
PF585002
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Email: cgabbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9799 row: d column: 01
High quality sequence stop: 604.
Location/Qualifiers
1..604
/organism="Mus musculus"

```

```

/strain="FVB/N"
/ab_xref="taxon:10090"
/clone="IMAGE:4218576"
/clone_1lb="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP library."

```

Query Match	14.6%	Score 465	DB 150	Length 604
Best Local Similarity	86.6%	Pred. No. 2.6e-96		
Matches 524	Conservative	0	Mismatches 80	Indels 1

OY	505	agcggaaaccccgagcccgacatcaacgltgaaagacagcaccggttcagcgcca	564
Db	1	AGTGGGCAACCCAGGGCCAGACATCATGTGTGATGAAGGATGACACGCTTGAACGCACTTA	60
OY	565	gaagcgcgcgtgagcccaagaagaagaatggaacatcgagacgttaagaactctgcgcgcgag	624
Db	61	GAGGCTAAGTGAACACAGAAAGAAAGTGGACCTGATGATTGAAGCACTGAAGCCTGAA	120
OY	625	gacacgcgcaaatacacactgcgcgcgtgtctgaacacgcgcgcgcgcacatcaacgcaccta	684
Db	121	GACATATGGCAATACACACGTCGGCTGATCTAACAAAGCGCGTGGCCATCAACGCCACACTAC	180
OY	685	aagtgatgaltgataccagcgcgaaccggttccaaagcccgctctcaacagcaacgcacccccg	744
Db	181	AAAGTGGATGATATCCAGCGGACCTGTTCCAAAGCCTGTCTCAACAGGACACACCCCTGTG	240

Db	241	AACACAACCGTGGACTTCGTTGGGAACAAGTCTTCCATGTGCAGAAGGTGGCAATTGACGTG	300
Oy	805	aagccgctgatccagtgtggtcgtgaacgcgctggagtaacgggccgaaggcgcccacaactcc	864
Db	301	AAGCCTGTATCTCAGTAGTGCTTGAAGCGGGTGGAACTACGGCTCCGAGGGACGCCCAACTCC	360
Oy	865	accatcatgatgtggtgcgccccagaagtttgtgtgtgtgtcccaagggtgaagtgtygttcgcg	924
Db	361	ACCATTGATGTGGGTGGCCAGAAAGTTGTGGTGTTCGCCAACGGGGTATGTGTGTACGG	420
Oy	925	cccgcaggtcttctaactcataaagctgctcatcacccgctgcgccccgaagaaacattgcygc	984
Db	421	CTGTAGTGTCTCTACTACTCAACAAGCTGTCTATCTCTCGGGCCCCGACGAGTAAGCTGGC	480
Oy	985	atgtacatctgtcctltgycgcaccaacacatgggtctaacagcttccgcagcgacctccaac	1044
Db	481	AGTACATCTGGCTTAGTGTGCAATAAATACCATAGGGCTACAGTTTCGGTAGCGCCTTCTCATT	540
Oy	1045	gtgcgtgcagaaccacaaccgcgcgaaggccaactgttgctctctcgtctcctgcgcacatagc	1104
Db	541	GTAATTACGACAGACCCCACAACTCCA-GGCCCTCATNTGGCTTCTTCATCTGTCATCCACAAGC	599
Oy	1105	ctggcc 1109	
Db	600	CTGGCC 604	

RESULT	38			
BF307860				
LOCUS	BF307860	981 bp	mRNA	EST
DEFINITION	601890670P1 NIH_MGC_17		Human sapiens	CDNA clone IMAGE:4131976 5',
				mRNA sequence.
ACCESSION	BF307860			
VERSION	BF307860.1	GI:11255017		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;			

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 981)	Mammalia: Euthera: Primates; Catarrhini; Hominoidea: Homo.	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .	National Institutes of Health, Mammalian Gene Collection (MGC)	
	Unpublished (1999)	Contact: Robert Strausberg, Ph.D.		
		<a href="mailto:robertstrauss@nih.gov">robertstrauss@nih.gov</a>		

cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at: [image.llnl.gov](http://image.llnl.gov)  
 Plate: LILCM030 row: k column: 17  
 High quality sequence step: 638.

FEATURES	SOURCE	LOCATION/Qualifiers
1..981		
/organism="Homo sapiens"		
/db_xref="taxon:9606"		
/clone="IMAGE:4131976"		
/clone_lib="NIH_MGC_17"		
/issue_type="rhabdomyosarcoma"		
/lab_host="DH10B (phage-resistant)"		
/note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using zap-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."		
208 a	311 c	289 g 172 t 1 others

[illegible]

Df	476	GGCACACCTGTGGGCGCTTCCGTCGTCACCTAGCCTCGGCCGCGGCATTTAGCTCGGCCGCGGTCGTCAAT	535
Qy	1125	cggcatcccaagcgcgagctgttcttcactctgttgacaccctgtctgtgtcgttcaggc	1184
Df	536	CGGCATCCACACCGGGGCTGCTTCCTAGCTTGCGGACCTGCTCTGTGTGTTCAAGGC	595
Qy	1185	ccag--aaagaagcgtlgcaaccccccgagctgtcccctcccctcctcctgttggaacgccgcgcg	1242
Df	596	CCAGCAAAGAAGACCCTGGTAGACCCCGCGCTGGCCCCCTCCCTGCTTGCGACCCGCGCGGTGG	655
Qy	1243	gggaagcggccgagccagccagcgagaacaagaccttcctctgtttggcgcgacctcaagcgt	1302
Df	656	AGGCCCCCGGACCGCGAG--GTGGAGAAAAGGCCCTTCCTCGTTTCGGGC--TAAGGCT	711
Qy	1303	ggccctgtgtgtggtgtgtgtgtgaagatgtgtgtcgcgcagcccc	1350
Df	712	GATCTGGGTGTGAGACTGCTGTGAGCGCATTTAGGTTCTGTGAGCCCCC	759
 RESULT_39			
LOCUS	BC387406	874 bp	mRNA EST
DEFINITION	602456172P1 NIH_MGC_15 Homo sapiens CDNA clone IMAGE:4584448 5', mRNA sequence.		
ACCESSION	BC387406		
VERSION	BC387406.1	GI:13280852	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	NIH-MGC http://mgi.nci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cga@b-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: NIH Intramural Sequencing Center Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LILCM1309 row: P column: 17 High quality sequence stop: 531. Location/Qualifiers		
FEATURES	Source		
	1..874		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4584448"		
	/clone_1lb="NIH MGC 15"		
	/tissue_type="adenocarcinoma cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAC(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"		
	BASE COUNT	225 a	287 c
	ORIGIN	253 g	109 t
 Query Match 14.3% Score 455.4; DB 153; Length 874; Best Local Similarity 96.8%; Pred. No. 4.4e-94; Matches 508; Conservative 0; Mismatches 11; Indels 6; Gaps 4			
Qy	902	caacgggtgagctgtgtgtgtgtgcggccgaaggtctcactcataaagcgtgcattacc	961
Df	2	CCACGGGTGACCTGTGTGTGTGTGCGACCGAGCGGTCTCACTCAATAAAGCTGTATACCC	61

Oy	962	gtgccgcgaagacatgctggagcatgatcattctgtcttggcccaaccatgtggctaca	1021
Db	62	GTGCCCGCCAGGACGATCGGGCATGTCATCTTGCCTTTGGGCCCAACACTTGCGCTACA	121
Oy	1022	gtcttcgcgaagcgctctctctcaacgtgtctgtgcagaagccaanaacgcgaaggccacctgttg	1081
Db	122	GCTTCCGACAGCGCCTTCTCACTCACCGTGTGCCAGACCAAAACCGCCAGGCGCACCTGTGG	181
Oy	1082	cctctctgtctctcgtgcacttagcctgcgcgttgcccgttgtatcgtggtatccccagccggcg	1141
Db	182	CCTCTCTGTCTCCCGGCCACTACCTGCGCGGTGG-CCGTGGTATGCGCATCCACGCGGCG	240
Oy	1142	ctgtcttatctcttggaaacctgtcctcgtgtgattttgcagaagcccaagaagccgtgtta	1201
Db	241	CTGTCTTATCTCTGGGGACCTGCTCTCTGTGGCTTTGGCAGAGCCCAAGAAGAACGCTGTGA	300
Oy	1202	ccccgcgcgtgcctccctccctcctgtcctgtgcagccgcgcgcggggaagccgcgcagcgca	1261
Db	301	CCCCCGCGCTTCCCTCCCTCCCGCTGCTGGGACACGCGCCCGCGGGAGCGS-CCGCGACCGCA	359
Oy	1262	gcggagaaaggaacctctccctcgtttgtgcgcgcgcctctaagcgtcgtgccttgttggggtgt	1321
Db	360	GGGAGAAAGAGACCTTCTCTGTGGGCGCCCTTAGCGGTGGCCCTGTGTGGGGCTGT	419
Oy	1322	gtgaagagcatgtgtctccgcgcagccccccag-cacttacctggggccagagccc--agt	1377
Db	420	GTGAGGACATGAGGCTCTCGGAGAACCCCAGAGACTTACTGTGGGCCCGAGACCCAAGTTT	479
Oy	1378	gtctgcacctaatgtgttaccaccaactctatacagaacatcacaca	1422
Db	480	GCTGAGCTAAGTTGTACCCCAAACCTTACAAAGACATTCACACA	524
RESULT	40		
BE780737			
LOCUS	BE780737	EST	20-OCT-2000
DEFINITION	601469609p1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3872547 5'		
ACCESSION	BE780737		
VERSION	BE780737.1	GI:10201935	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 951)		
JOURNAL	NIH-MGC http://mgc.ncl.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-re@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <a href="http://image.lnl.nih.gov">http://image.lnl.nih.gov</a> Plate: LHAM9627 row: b column: 04 High quality sequence stop: 660. Location/Qualifiers 1..951 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3872547" /clone_id="NIH_MGC_67" /tissue_type="retinoblastoma" /lab_host="DH10B (phage-resistant)" /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."		

BASE COUNT 185 a 328 c 297 g 141 t  
 ORIGIN

Query Match 14.2%; Score 453.6; DB 140; Length 951;  
 Best Local Similarity 98.1%; Pred. No. 1.2e-93;  
 Matches 512; Conservative 0; Mismatches 4; Indels 6; Gaps 5;

OY 909 tgaactgtgtgtcgccgacgagctcctactactaataagctgtcatcaccgtgtccg 968  
 |||||||  
 Db 1 TGACGTGTGTGCGCG -CCGAGCGCTCTACTCAATAGCTGCTCATCCCGTGGCCG 59  
 |||||||

OY 969 ccagagcagatgaggcagatgacatctgctgtggcccaacacatgagctaaagcttcg 1028  
 |||||||  
 Db 60 CCAGAGCGATGCGGCGATGCTATCTGCTTGGCCGCAACACCATAGGCTTACAGCTTCCG 119  
 |||||||

OY 1029 cagcgcccttcacacgltgctcagacccaaacacgcccagggccacactgtgtgctctc 1088  
 |||||||  
 Db 120 CAGCGCTTCTCAGCGTCTGCTGCAAGCCAAACCGCAGGCGCACCTGTGCTCTC 179  
 |||||||

OY 1089 gtctctgacactagactgtgcgtgtgcgtgtgtcatcgtgacccagccggtgtctt 1148  
 |||||||  
 Db 180 GTCTCTGCGGCGACTAGCTGCTGCGCTGCTGTCAATCGCATCCAGCGGCGTGTCTT 239  
 |||||||

OY 1149 catcctgggacacccctgctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1208  
 |||||||  
 Db 240 CATCTGCGGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 299  
 |||||||

OY 1209 gctctgcccctccctgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1268  
 |||||||  
 Db 300 GCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 357  
 |||||||

OY 1269 caagagccttcctgt 1328  
 |||||||  
 Db 358 CAAGGACCTTCTCTGCTGTGCGG -CCTCAGCGCTG -CCTGAGTGTGGCTGTGTAGGA 415  
 |||||||

OY 1329 gcatgtgtctcgc 1388  
 |||||||  
 Db 416 GCATGTGTCTCGCGGCGCGCCCGCAGCACTTACTGTGGCGCCAGGCGCTGTGTG -CCTAA 474  
 |||||||

OY 1389 gtgtgtaccccaactctaacagacatccacacacacacacacacacacacacacac 1430  
 |||||||  
 Db 475 GTGTACCCCAACTCTACACAGACATCCACACACACACACACACACACACACACAC 516  
 |||||||

RESULT 41  
 AI921390 456 bp mRNA EST 08-MAR-2000  
 LOCUS WO24C03.x1 NCI\_CGAP\_Pan1 Homo sapiens cDNA clone IMAGE:2456260 3'  
 DEFINITION similar to contains Alu repetitive element; mRNA sequence.  
 ACCESSION AI921390  
 VERSION AI921390.1 GI:5657354  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 456)  
 TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-r@mail.nih.gov  
 Life Technologies catalog #: 11548-013  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www.bio.llnl.gov/bbrp/image/image.html  
 Insert length: 1249 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 424.  
 Location/Qualifiers

source 1..456  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2456260"  
 /issue\_type="NCI\_CGAP\_Pan1"  
 /tissue\_type="adenocarcinoma"  
 /lab\_host="DH10B"  
 /note="Organ: pancreas; Vector: pCMV-Sport6; Site: 1; Salt;  
 Site: 2; Note: Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"

BASE COUNT 113 a 113 c 133 g 97 t  
 ORIGIN

Query Match 13.9%; Score 444; DB 103; Length 456;  
 Best Local Similarity 99.8%; Pred. No. 1.7e-91;  
 Matches 455; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 2730 ctccatcccgctctgtccctgtgctgtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2789  
 |||||||  
 Db 456 CTCATCCCGCGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 397  
 |||||||

OY 2790 ccagagatccctactgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2849  
 |||||||  
 Db 396 CCAGGAGTCCCTACTGCTGTGGCTGTGGGCTGTGGGCGACAGCAGCCCAAGCTGAGAG 337  
 |||||||

OY 2850 gctgagacccatgt 2909  
 |||||||  
 Db 336 GCTGAGCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 277  
 |||||||

OY 2910 gctctgtattatatttaagaanaatgaagataatattatattatattatattatatt 2969  
 |||||||  
 Db 276 GCCTTGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 217  
 |||||||

OY 2970 gtgtgagagactgt 3029  
 |||||||  
 Db 216 GTTGAAGGAGCTGT 157  
 |||||||

OY 3030 atgaccacaccccgctcagcagacaga -caaccccccaacccactgtgtgtgtgtgtgt 3088  
 |||||||  
 Db 156 ATGACACACACCCCGTCCAGGCGACAGACACACCCCGCAGCTGTGTGTGTGTGTGTGT 97  
 |||||||

OY 3089 gatctctgaattatattatattatattatattatattatattatattatattatatt 3148  
 |||||||  
 Db 96 GATCTCTGTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 37  
 |||||||

OY 3149 aaacatgaagatgacatccttcctcccaaaaaaaa 3184  
 |||||||  
 Db 36 AAACATGAAGATGATCTTCTCTCCCAAAAAAAA 1  
 |||||||

RESULT 42  
 AM083114 443 bp mRNA EST 14-OCT-1999  
 LOCUS Xc06b10.x1 NCI\_CGAP\_CO21 Homo sapiens cDNA clone IMAGE:2583451 3'  
 DEFINITION similar to contains Alu repetitive element; contains element MSN1  
 repetitive element; mRNA sequence.  
 ACCESSION AM083114  
 VERSION AM083114.1 GI:6038266  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 443)  
 TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-r@mail.nih.gov  
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael





	Db	127	TGTAGAGTTTACGTCGAAGCCCCCGAATATTAAATTATTGGTAACAATGAAGTGCAAT	68
OY	3165	ccttccctcccaaaaaaaaaa	3186	
Db	67	CCTTCCCTCCCAAAAAAAAAAAAA	46	
	RESULT 44			
	BP969458	BF969458	EST	22-JAN-2001
	LOCUS	60227175.f1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4360068 5'		
	DEFINITION	mRNA sequence.		
	ACCSSION	BF969458		
	VERSION	BF969458.1 GI:12336673		
	KEYWORDS	EST.		
	SOURCE	human.		
	ORGANISM	Homo sapiens		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
	REFERENCE	1 (bases 1 to 944)		
	AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.		
	TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
	JOURNAL	Unpublished (1999)		
	COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs.rem@nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Plate: LLM10000 row: k column: 13 High quality sequence stop: 592. Location/Qualifiers		
	FEATURES			
	Source	1..944		
		/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
		/clone="IMAGE:4360068"		
		/clone_lib="NIH_MGC_84"		
		/tissue_type="adrenal cortex carcinoma, cell line"		
		/lab_host="DH10B (phage-resistant)"		
		/note="Organ: adrenal gland; Vector: PCMV-SPOrt6; Site:_1; Note: Site_2: Salt; Cloned unidirectionally. oligo-dT primed. Average insert size 1.229 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."		
	BASE COUNT	277 a 284 c 228 g 155 t		
	ORIGIN			
	Query Match	13.7%:	Score 437:	DB 172: Length 944:
	Best Local Similarity	86.4%:	Pred. No. 7.6e-90:	
	Matches 649: Conservative	0: Mismatches 75: Indels 27: Gaps 14:		
OY	1796	caacactactgatatgatattatgcacaacaaatgaagcgacacagtgtcccttaaggcaca	1855	
Db	1	CACACTGCCCTGGATGTGATGTATGCACACACATGCCGCCACAGCTGCTCCTTAAGGCACA	60	
OY	1856	cgtagcacacagaagcatatgcacagataltcgccgttgggcaca -cagataagctgccaa	1914	
Db	61	CGTAGCACACAGACGACCATGTGCGAACATAAGACACTGCTGCTGAAGCCCAA	120	
OY	1915	atgcacgcacacagcagagacctgcacagaccttaaaggacatttgtctctaagctata	1974	
Db	121	ATGCACGCACACCCACAGAAGACATGCCAACGAATTAAGAGCACTGCTGCTGAACATACA	180	
OY	1975	cagcagacaccatgcgcagatgtgcctgcctggacacacacacacacagatattgtctgc	2034	
Db	181	CAGCAGACCCCAATGCGCAGATGTGCTGCGCTGGACACACACACACAGCATATGCTGTC	240	
OY	2035	tggacgcacacacacgtgcagatatgttatccc-ggacacacacac-gttcacagatatgtctgc	2092	

Db	241	TGGACGCACACAGCTGGAGATATGATCTCTGGACACACACTGTGCACAGATATGCTGGC	300
QY	2093	ttgacacacagataatgtctgtcc--ttgacacacacatgcaacgga--catgtccttgagaccac	2149
Db	301	TGGACACACACATATATCTGCTGCATATGGACACACACATGACACGGAAATATTTGCCGTGGACACA	360
QY	2150	cacacacacagatgttggacagatatgtatgtct-----tgacacgcacacacatgagat	2203
Db	361	CACACACACACAGTGTGGACAGATATGTCTGTCTTGGACAGAGCACACACATTTGATATTA	420
QY	2204	atgtctccttgacacacacacttccagaca--cacgtgcacagcgcgagata--tgtctgtc	2260
Db	421	TGGCGCTCTGGACACACACTTCCAGACATCATCTGTCACAGCGCAGATATGCTCTGCTTG	480
QY	2261	gacacacgcgagatatgtctgtctagttaacacacacacacgcaacatgtctgtctcgagaca--ca	2318
Db	481	GACACACGCACATATATGCTGTCTTAATCTACACACACACACCGGACACATGCTGTCCGACACACAT	540
QY	2319	cacacagcatgacagatatgtctgtctccgagaca--cacacacgacacagatatgtctgtc	2376
Db	541	CACACGCATGGACAGATATGTCTGTCTCCGACATTCAGCACACGACAGCAGATATGCTGTGG	600
QY	2377	gacacacacacagataatgtctgtcctcaacacttcaacacacgctgcagatatattgtcctgacac	2436
Db	601	ATACACGCAACAAGATAGTGTGCTGCTCAACACTCA--AGAAGTGCAGGATATATGCTGGACAC	659
QY	2437	acacatgtgcacagatatgtctgtcttgagacatagcacacacgcttgagatatgtctgtccgat	2496
Db	660	ACACGTTCACCGAATATGCTGTCTGGACGTCG---CCCGTGCAGAT--TGTGTGCG---712	
QY	2497	acacacgcacacacatgtcagatatgtctg	2527
Db	713	ATACTAAGCAGACACATGCAATTTGGTGC	743

```

RESULT 45
LOCUS AA442297
DEFINITION AA442297 441 bp mRNA EST 02-JUN-1997
v554c08.t1 Soares_testis_NHT Homo sapiens CDNA clone IMAGE:757454
5' similar to contains element LTR4 repetitive element ; , mRNA
sequence.
ACCESSION AA442297
VERSION AA442297
KEYWORDS AA442297.1 GI:2154175
SOURCE EST.
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 441)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Giesel,G., Jost,S.,
Kucab,T., Lacy,M., Le,N., Lennon,G., Mairr,M., Martin,T., Moore,B.,
Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wyllie
,T., Waterson,R. and Wilson,R.
Washu-Merck EST Project 1997
Unpublished (1997)
TITLE CONTACT: Wilson RK
JOURNAL Washington University School of Medicine
COMMENT 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 398.
FEATURES
Source
Location/Qualifiers
1..441
/organism="Homo sapiens"
/db_xref="GDB:5978068"
/db_xref="taxon:9606"
/clone="IMAGE:757454"
/clone_lib="Soares_testis_NHT"

```

```
/sex="male"  
/lab_host="DH108"  
/note="vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA  
was prepared from mRNA obtained from Clontech Laboratories  
, Inc., and primed with a Not I - oligo(dT) primer (5'  
TGTTCACATCTGAGTGGAGCGGCCGCAATTTTTTTTTT 3').  
Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization to Col5, and was  
constructed by Bento Soares and M. Fatima Bonaldo."
```

```
BASE COUNT      133 a      143 c      97 g      68 t  
ORIGIN
```

```
Query Match      13.7%; Score 436.2; DB 7; Length 441;  
Best Local Similarity 99.3%; Pred. No. 1e-89;  
Matches 438; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1922 caccgcacagagacatgcccagaacatatacagatgctgctgaacatatacagcgcac 1981  
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 1 CACACGCACAGACATGCCAACAATACAGACATGCTGCTGAACATACACACGCAC 60  
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1982 acccatggcgagatgctgctgctgacacacacacacagatatagtctgtgagcg 2041  
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 61 ACCCATGGCGAGATGCTGCTGGACACACACACACAGGATATGCTGTGGACGC 120  
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2042 acacacgtgcagatatagtatccggacacacacgtgcacagatatagtctggaacac 2101  
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 121 ACACACGTGCAGATATGTTCCGGACACACACGTGCACAGATATGCTGCTGGACACAC 180  
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2102 agataatgctgcttgacacacacatgacgatatgtccttgacacacacacacacg 2161  
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 181 AGATAATGCTGCTTGACACACACATGCAGATATGCTGCTGGACACACACACACG 240  
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2162 tgtgcacagatagctgctgctgacacacacacacatgacatatgtccttgacacaca 2221  
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 241 CGTGACAGATATGCTGTGTGGACAGGCACACATGCAGATATGCTGCTGGACACACA 300  
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2222 ctccagacacacgtgcacagcgacagatatgtctgcccctggacacacgagatatgtctg 2281  
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 301 CTTCAGACACACAGTGCACAGGCGCAGATATGCTGCTGGACACACCGCAGATATGCTGTC 360  
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2282 tagtcacacacacacagacatgctgttccggacacacacacacacgcatgacagatalgtctg 2341  
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 361 TAGTCACACACACAGCAGCATGCTGTCCGGACACACACACGCGATGCACAGATATGCTG 420  
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 2342 tcggagacacacacagcagcg 2362  
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 421 TCCGGACACACACACGCGACGC 441
```

```
Search completed: August 17, 2001, 03:05:12  
Job time: 2578 sec
```

**THIS PAGE IS BLANK**

OM of: US-09-598-042A-2 to: SwissProt\_39: \* out\_format : pfs

Date: Aug 17, 2001 2:19 AM

About: Results were produced by the Gencore software, version 4.5  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```

-MODEL=lrtime,n2p,model -DEV=xlp
-O=cgml2,1/USPTO.spool/US09598042/rnat,150822001,141544_26384/app-query.fasta_1.3297
-DB-SwIsnpProt_39 -OFMT=fastan -SUPERF=rrp -GAPOP=12.000
-GAPEXT=0.000 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=dlsum62
-TRANS=human40.cdi -LIST=45 -DCCALC=GENE200 -THR_SCORE=Dct
-THR_MAX=100 -THR_MIN=0 -ALIGN=45 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MILEN=0 -MAXLEN=2000000000
-USER=US09598042.ecgml_1_57 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

```

```
Search information block:
Query: US-09-598-042A-2
Query length: 3186
Database: SwissProt_39:*
Database sequences: 93435
Database length: 34255486
Search time (sec): 77.360000
```

Sequence	Strid	Orig	ZScore	EScore	Len	Document
SwissProt_39:FER3_MOUSE	+	470.00	393.57	3.5e-15	801	061851 mus musculus (mouse)
SwissProt_39:FER3_HUMAN	+	490.00	383.27	1.3e-14	806	061881 mus musculus (mouse)
SwissProt_39:FER3_XENLA	+	469.00	376.91	2.9e-14	812	P22607 xenopus laevis (Xenopus laevis)
SwissProt_39:CEK3_CHICK	+	463.00	372.08	5.3e-14	823	P18461 gallus gallus (chicken)
SwissProt_39:FER2_HUMAN	+	462.00	371.31	5.9e-14	821	P21802 homo sapiens (human)
SwissProt_39:FER2_MOUSE	+	461.00	370.52	6.5e-14	821	P21845 mus musculus (mouse)
SwissProt_39:FER4_HUMAN	+	459.00	369.50	7.6e-14	802	P22455 homo sapiens (human)
SwissProt_39:CEK2_CHICK	+	456.50	367.10	1.0e-13	806	P18460 gallus gallus (chicken)
SwissProt_39:FER1_CHICK	+	442.00	355.64	4.5e-13	808	003142 mus musculus (mouse)
SwissProt_39:FER1_MOUSE	+	428.50	344.87	7.4e-13	819	P21804 gallus gallus (chicken)
SwissProt_39:FER2_XENLA	+	425.00	343.57	1.7e-12	822	P16092 mus musculus (mouse)
SwissProt_39:FER1_HUMAN	+	423.50	340.92	2.9e-12	822	003564 xenopus laevis (Xenopus laevis)
SwissProt_39:FER1_RAT	+	423.50	340.92	2.9e-12	822	P1362 homo sapiens (human)
SwissProt_39:PER2_HUMAN	+	379.50	307.76	2.6e-10	654	001742 homo sapiens (human)
SwissProt_39:NTC3_MOUSE	+	313.50	247.04	1.7e-07	2318	061882 mus musculus (mouse)
SwissProt_39:PEBM_HUMAN	+	312.00	241.50	1.9e-07	4393	P98180 homo sapiens (human)
SwissProt_39:PPD_HUMAN	+	301.50	238.89	6.0e-07	1912	P23468 homo sapiens (human)
SwissProt_39:LAR_DROME	+	299.00	236.51	7.7e-07	2029	P16621 drosophila melanogaster (fruit fly)
SwissProt_39:DC2_MOUSE	+	298.00	238.03	8.9e-07	1447	P70211 mus musculus (mouse)
SwissProt_39:BAR3_CHITE	+	297.50	236.53	9.1e-07	1700	003376 chironomus tentans (midge)
SwissProt_39:PEBM_MOUSE	+	293.00	227.66	1.3e-06	4707	005793 mus musculus (mouse)
SwissProt_39:TEKX_HUMAN	+	293.00	226.67	1.3e-06	4289	P22105 homo sapiens (human)
SwissProt_39:FER2_DROME	+	292.50	233.86	1.6e-06	51052	009147 drosophila melanogaster (fruit fly)
SwissProt_39:DCQ_HUMAN	+	292.50	233.69	1.5e-06	14447	P43146 homo sapiens (human)
SwissProt_39:NTC4_MOUSE	+	289.00	228.84	2.1e-06	1964	P31695 mus musculus (mouse)
SwissProt_39:MOC1_HUMAN	+	286.50	229.92	2.9e-06	1255	P15941 h. mucin 1 precursor
SwissProt_39:NCB2_HUMAN	+	284.50	231.76	3.8e-06	761	P13552 homo sapiens (human)
SwissProt_39:YHL1_EBV	+	283.00	231.55	4.5e-06	660	P03181 Epstein-Barr virus (strain 95-1)
SwissProt_39:NRG_DROME	+	274.00	220.15	1.0e-05	1239	P02041 drosophila melanogaster (fruit fly)
SwissProt_39:FER1_DROME	+	273.50	223.36	1.2e-05	739	P07407 drosophila melanogaster (fruit fly)
SwissProt_39:PK7_HUMAN	+	272.00	219.57	1.3e-05	1070	P13308 homo sapiens (human)
SwissProt_39:KMT5_HUMAN	+	267.50	212.05	1.9e-05	1914	015746 homo sapiens (human)
SwissProt_39:AKO1_HUMAN	+	266.00	215.03	2.4e-05	1040	002246 homo sapiens (human)
SwissProt_39:PEPE_HUMAN	+	264.00	209.33	2.7e-05	1897	P10586 homo sapiens (human)
SwissProt_39:AKO1_RAT	+	263.00	211.87	3.5e-05	1040	P22065 rattus norvegicus (brown rat)
SwissProt_39:PK7_CHICK	+	261.50	211.41	3.7e-05	1051	009104 gallus gallus (chicken)
SwissProt_39:NCAL_BOVIN	+	259.50	211.25	4.7e-05	853	P18363 bos taurus (bovine)
SwissProt_39:MOC1_HUMAN	+	259.00	198.55	3.9e-05	5179	002817 homo sapiens (human)
SwissProt_39:NCAL_RAT	+	257.00	209.24	6.0e-05	858	P13596 rattus norvegicus (brown rat)

seq_name	Swissprot_39	FCGR3_MOUSE
SwissProt_39:CD14_HUMAN	256.50	208.93
SwissProt_39:EB15_CAREL	256.50	207.53
SwissProt_39:CONT_MOUSE	255.50	206.88
SwissProt_39:RNC1_RAT	253.00	198.70
SwissProt_39:CDML_MOUSE	252.50	203.07
		9.0e-05
		1260
		PI1391 homo sapiens (human)
		PI0656 caenorhabditis el
		PI13860 mus musculus (mouse)
		PI07008 rattus norvegicus (mouse)
		PI1627 mus musculus (mouse)

```
seq_documentation_block:
ID      FGR3_MOUSE          STANDARD:          PRT:          801 AA.
AC      Q61851; O63834;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      15-JUL-1999 (Rel. 38, Last annotation update)
DE      FIBROBLAST GROWTH FACTOR RECEPTOR 3 PRECURSOR (FGR3-3) (EC 2.7.1.112)
EN      (HEPARIN-BINDING GROWTH FACTOR RECEPTOR).
GN      FGR3 OR MFR3 OR SAM3.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX      NCBI_TaxId=10090;
[1]
RN      R1
RP      SEQUENCE FROM N.A.
RM      MEDLINE=92355591; PubMed=1379594;
RA      Ornitz D.M., Leder P.;
RT      "Ligand specificity and heparin dependence of fibroblast growth
RT      factor receptors 1 and 3."
RL      J Biol. Chem. 267:16305-16311(1992).
[2]
RN      R2
RP      SEQUENCE FROM N.A.
RM      TISSUE=Brain;
RX      MEDLINE=93177694; PubMed=8382556;
RA      Katoch O., Hattori Y., Sasaki H., Sakamoto H., Fujimoto K.,
RA      Fujii T., Sugimura T., Terada M.;
RT      "Isolation of the complementary DNA encoding a mouse heparin-binding
RT      growth factor receptor with the use of a unique Kinase Insert
RT      sequence."
RL      Cancer Res. 53:1136-1141(1993).
CC      CC
CC      -1- FUNCTION: RECEPTOR FOR ACIDIC AND BASIC FIBROBLAST GROWTH
CC      FACTORS. DIFFERENTIALLY BINDS ACIDIC PGF.
CC      CC
CC      -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC      CC
CC      -1- TISSUE SPECIFICITY: IN EMBRYO, EXPRESSED IN HEART, LUNG, KIDNEY,
CC      SKIN, HEAD AND LIVER BUT NOT IN MUSCLE. IN ADULT, HIGHEST LEVELS
CC      IN BRAIN. ALSO EXPRESSED IN LIVER, LUNG, KIDNEY, TESTIS, OVARY
CC      AND UTERUS. VERY LOW LEVELS IN HEART, THYMUS, SPLEEN AND MUSCLE.
CC      CC
CC      -1- DEVELOPMENTAL STAGE: EXPRESSED IN EMBRYO FROM MID-GESTATION AND
CC      IN ADULT.
CC      CC
CC      -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
CC      FAMILY.
CC      CC
CC      -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE DOMAINS.
-----
CC      CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation at
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
CC      CC
EMBL: MB1342; AAA39535.1; -.
DR      EMBL; S56291; AAB25535.1; -.
DR      HSSP; P11362; IFGI.
DR      MGd; MG1:95524; FGfr3.
DR      InterPro; IPR000719; -.
DR      InterPro; IPR001245; -.
DR      InterPro; IPR003006; -.
DR      Pfam; PF00047; Ig_3.
DR      Pfam; PF00069; pkinase; 1.
DR      PRINTS; PR00109; TYRKINASE.
DR      PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR      PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR      PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
FM      Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;
```

140 glnsprhrnglvalaProtyttr.....Tharprprogl 151

441 CAAATGATAGGCGCGGGTATCGACAGCGCGCGGTGGAGTACGTCGGTGGCG 450

151 uarMetasrlyuLysleuLeuAlaValProAlaAlaAsnthValAlaThr 168

491 TCAAGTGGGTGGCGAGCGGGACGCGCGTGGCGCGCATCGATCGATGGATGAG 540

168 hehYcysProAlaAlaIleGlyAsnProThrProserIleSerThrLeuLeu 184

541 GACGACGAGCGCGTTGACG.....CGCCAGAGGCGCGCTGAGCCGAC 581

185 AsnGlyuSerGluPheArgGlyGluHisArgIleGlyIleLeuLysLeuAla 201

582 GAACAGACAGTGGACACTGAGACCTGAAGAACCTCGCGCGCGAGACAGCG 631

201 gHisGlnIleThrPheLeuValMetGluSerValAlaProSerSerArg 218

632 GCAATATACCTCGCGCGGTGGTGCAGACCGCGCGGGCGCGCATCGACCGAC 681

218 LysAsnTyThrCysValAlaGluLysnLysPheGlySerIleArgGlnThr 234

682 TACAGGTGATGTGATCGACAGCGACCGCGTTCCAAAGCGCGTGGTACAGG 731

235 TyrThrLeuAspValIleGluIArgSerProHisArgProIleLeuGlnAl 251

732 CACGACCGCGCGTGAACACAGACGGTGGACTTCGGGGGGACACAGTCCTTC 781

251 agLYeuProAlaAsnGlnThrAlaIleLeuGlySerAspAlaIlePheH 268

782 AGTGCAGCGTCCGAGACGACGTGAAGCGCGGATGATCGATGGCTGACAGCG 831

268 TSCYLSValTYLysSerAspAlaGlnProHisIleGlnThrLeuLysHis 284

832 GTGAGTACGCGCGCGGAGCGCGGACACATCCACATCGATGGTGGCGG 881

285 ValGluVal.....AsnGlySerLysValGlyProAspArgI 296

882 CCAAGAGTTTGTG...GTGCTGCCACCGGCTGAGCTGTGGTGGCGCGCG 928

296 yThrProTyAlaThrValLeuLysThrAlaGly.....AlaAsnthrT 311

929 ACGGCTCTACCTCAATATAGTGGTGCATACCGCGCGTGGCGCGACAGCAT 978

311 hrAspLysGluLeuGlnValLeuSerLeuHisAsnValAlaThrPheGluAsp 327

979 GCGGAGTATACATCTGCGTGGCGGCGCAACACATAGGCTAGCTGCGCG 1028

328 AlaGlyuTyuTyThrCysLeuAlaGlyLysnSerIleGlyPheSerGlnHis 344

1029 CAGCGCGCTTCACACCGTGGTGGCA...GACCCAAACCGCGGACGGCGAC 1075

344 sSerIleThrPheLeuValValAlaLeuProAlaGlnGluGluMetGluThr 361

1076 CTGTGCGCTCCGTCGTCGCGGACACTAGCTGCGCGTGGCGCGGTGATCTC 1125

361 spGluAlaIleGlySerValTyuAlaGlyValLeuSerTyGlyValAlaPhe 377

1126 GGCATCCAGCGCGGCGGTCTTCATCTCTGGGACACCTGTCCTGTGGCT 1175

378 .....PheLeuPheIleLeuValValAlaAlaValIleLe 389

1176 TTGGCAGCGCGCAAGAAAGACCGGTGACACCGCGCGCTGCGCGTCCGCG 1224

389 uCysArgLeuArgSerProProLysGlyLeuGlySerProThrAlaH 406

1225 .....CCTGGGACCGCGCGCGGAGACGGCGCGGAC 1257

406 IsLYuValSerArgPheProLeuLysArgGlnValSerLeuGluSerAsn 422

1258 CGCAGCGGACAAAGACCTTCC.....TCGTGGCGCGCTCGACGCG 1301

423 SerSerMetAsnSerAsnthProLeuValArgIleAlaArgLeuSerSe 439







```

CC or send an email to license@slsb.sib.ch).
-----
DR EMBL, U24491, AAA86868.1, -.
DR PIR, A36477, A36477.
DR HSSP, P1362, IEG1.
DR InterPro, IPR000719, -.
DR InterPro, IPR001245, -.
DR InterPro, IPR003006, -.
DR Pfam, PF00047, Ig, 3.
DR Pfam, PF00069, pkinase, 1.
DR PRINTS, PR00109, TYRKINASE.
DR PROSITE, PS00107, PROTEIN_KINASE_ATP, 1.
DR PROSITE, PS00109, PROTEIN_KINASE_TYR, 1.
DR PROSITE, PS50011, PROTEIN_KINASE_DOM, 1.
DR RecPfort, Glycoprotein, Tyrosine-protein kinase; ATP-binding;
KW Transferrase; Phosphorylation; Transmembrane; Immunoglobulin domain;
KW Signal.
FT SIGNAL 1 20
FT CHAIN 21 812 BY SIMILARITY
FT DOMAIN 21 371 FIBROBLAST GROWTH FACTOR RECEPTOR 1.
FT TRANSMEM 372 393 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 394 812 POTENTIAL.
FT DOMAIN 472 812 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 472 761 PROTEIN KINASE.
FT DOMAIN 47 107 IG-LIKE DOMAIN.
FT DOMAIN 167 233 IG-LIKE DOMAIN.
FT DOMAIN 266 344 IG-LIKE DOMAIN.
FT NP_BIND 478 486 ATP (BY SIMILARITY).
FT BINDING 508 508 ATP (BY SIMILARITY).
FT ACT_SITE 617 617 BY SIMILARITY.
FT MOD_RES 648 648 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 326 326 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 812 AA; 90502 MW; B06333BAFEAC5C9B CRC64;

alignment_scores:
Quality: 469.00 Length: 437
Ratio: 1.757 Gaps: 15
Percent Similarity: 61.098 Percent Identity: 30.206

alignment_block:
US-09-598-042A-2 x FGRL_XENLA .
Align seg 1/1 to: FGRL_XENLA from: 1 to: 812

103 GCCCGCGCTGGCCGCACATATGCGGCTGCAGTCCCAAGTGGAGGGGACC 152
||||| ||| ::| ||||||| :::::: |||
44 AAGAGnProGlyAspArgIleThrLeuGlnGlyArgLeuArgGlnAspVal 60
153 GCCGGCGCGTGAACCATGTGGACCAAGAGTGGCGGCGCAACCATCCACAGCGGT 202
::::: |||:::||||:||||| :::: |||
60 IGIserIleasn...TTPValIysasnGlyValaGlnLeu...SegIUT 75

203 GGAGCCGCTTCCGCGTGCATGCGCGCAGAGGGCTCAAGGTGAACGAGGTGGAG 252
::::| |||::: ::| ::::::: :::::::
75 hrAsnArghrArghIleThrIleArgIleGlnIleGlnIleSerAsnIaGly 91

253 CGGAGAGATGCCGCGGTGATGTCGTGCAAGGCCACCAACGCGCTTGGCGAG 302
||||| |||||||:::|||| ::::::| |||
92 ProGluAspAsnGlyValItyrAlaGly...ValThrAsnGly..... 104

303 CCTGAGCGTCACTACACCCCTC.....GTGCGTGCATGATGCATTATA 343
||| ::::| |||||::: ||| ::::::| |||
105 .ProSerArghrItyrThrValIleuGlySerValAsnValSerAspAlaL 121

```

[illegible]

```

406  erg1nleuAlaValHisLysleuAlaAlaLysSerIleProValArgArgIln 422
1240  CCGGGAGCGCCCGCGACCGCAGCGACAGACAGACGCTTCCCTCGTTGGC 1289
443  ValTrnValSerGlyAspSerSerSer..... 431
1290  CGCCCTCAGCGCGCTGGCCGCTGGGTGGGCGCTGTGTGAGAGCATGGGCTTC 1339
442  .....SerMetAsnSerGlyValIleleuValAlaArgArgLeuSerSerS 446
1340  CGGCAGCCCCC 1350
446  erg1YThrPro 449

seq_name: SwissProt_39:CEK3_CHICK

seq_documentation_block:
ID  CEK3_CHICK  STANDARD;  PRT;  823  AA.
AC  P18461;
DT  01-NOV-1990 (Rel. 16, Created)
DT  01-NOV-1990 (Rel. 16, Last sequence update)
DT  01-NOV-1995 (Rel. 32, Last annotation update)
DE  TYROSINE KINASE RECEPTOR CEK3 PRECURSOR (EC 2.7.1.112).
GN  CEK3.
OS  Gallus gallus (Chicken).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
CC  Gallus.
CC  NCBI_TaxID=9031;
OK  [1]
RN  SEQUENCE FROM N.A.
RP  MEDLINE=90332672; PubMed=2165604;
RA  Pasquale E.B.;
RT  "A distinctive family of embryonic protein-tyrosine kinase
RT  receptors."
RL  Proc. Natl. Acad. Sci. U.S.A. 87:5812-5816(1990).
CC  -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN
CC  -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
CC  FAMILY.
CC  -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE DOMAINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: M35196; AAA48665.1; -.
DR  PIR: B35963; B35963.
DR  HSSP: P1362; IFS1.
DR  InterPro: IPR000719; -.
DR  InterPro: IPR001245; -.
DR  InterPro: IPR003006; -.
DR  Pfam: PF00047; Ig: 3.
DR  Pfam: PF00069; pkinase: 1.
DR  PRINTS: PRO0109; TYRKINASE.
DR  PROSITE: PS00107; PROTEIN_KINASE_ATP_1.
DR  PROSITE: PS00109; PROTEIN_KINASE_TYR_1.
DR  PROSITE: PS50011; PROTEIN_KINASE_DOM_1.
KW  Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;
KW  transferase; Phosphorylation; Transmembrane; Immunoglobulin domain;
KW  Signal.
KW  SIGNAL.
FT  CHAIN 1 23 POTENTIAL.
FT  CHAIN 24 823 TYROSINE KINASE RECEPTOR CEK3.
FT  TRASNEM 24 379 EXTRACELLULAR (POTENTIAL).
FT  DOMAIN 380 400 POTENTIAL.
FT  DOMAIN 401 823 CYTOPLASMIC (POTENTIAL).
FT  DOMAIN 58 117 IG-LIKE DOMAIN.
FT  DOMAIN 174 240 IG-LIKE DOMAIN.
FT  DOMAIN 273 351 IG-LIKE DOMAIN.

```



ID FGR2\_HUMAN STANDARD: PRT: 821 AA.  
 AC P21802; P18443;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE FIBROBLAST GROWTH FACTOR RECEPTOR 2 PRECURSOR (FGR2-2) (EC 2.7.1.112)  
 DE (KEATINOCYTE GROWTH FACTOR RECEPTOR).  
 GN FGR2 OR BEK OR BFR-1 OR KSAM-1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Neonatal brain stem;  
 RX MEDLINE=90360977; PubMed=1697263;  
 RA Dione C.A., Crumley G.R., Bellot F., Kaplow J.M., Searfoss G.,  
 RA Ruta M., Burgess W.H., Jaye M., Schlessinger J.;  
 RT "Cloning and expression of two distinct high-affinity receptors  
 RT cross-reacting with acidic and basic fibroblast growth factors";  
 RL EMBO J. 9:2685-2692(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91045961; PubMed=2172978;  
 RA Housaint E., Blanquet P.R., Champion-Arnaud P., Gesnel M.C.,  
 RA Torriglia A., Coutois Y., Brethnach R.;  
 RT "Related fibroblast growth factor receptor genes exist in the human  
 RT genome";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8180-8184(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91274356; PubMed=1647213;  
 RA Seno M., Sasada R., Watanabe T., Ishimaru K., Igarashi K.;  
 RT "Two cDNAs encoding novel human fgr receptor";  
 RL Biochim. Biophys. Acta 1089:244-246(1991).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Stomach cancer;  
 RX MEDLINE=90332706; PubMed=2377625;  
 RA Hattori Y., Odagiri H., Nakatani H., Miyagawa K., Naio K.,  
 RA Sakamoto H., Katoh O., Yoshida T., Sugimura T., Terada M.;  
 RT "K-sam, an amplified gene in stomach cancer, is a member of the  
 RT heparin-binding growth factor receptor genes";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:5983-5987(1990).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92212948; PubMed=1313574;  
 RA Katoh M., Hattori Y., Sasaki H., Tanaka M., Sugano K., Yazaki Y.,  
 RA Sugimura T., Terada M.;  
 RT "K-sam gene encodes secreted as well as transmembrane receptor  
 RT tyrosine kinase";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2960-2964(1992).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=93016048; PubMed=1400433;  
 RA Dell K.R., Williams L.T.;  
 RT "A novel form of fibroblast growth factor receptor 2. Alternative  
 RT splicing of the third immunoglobulin-like domain confers ligand  
 RT binding specificity";  
 RL J. Biol. Chem. 267:21225-21229(1992).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RA Steinberger D., Mueller U.;  
 RT Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.  
 RN [8]  
 RP VARIANTS CS HIS-340; ARG-342; SER-342; TYR-342 AND CYS-354.  
 RX MEDLINE=95078932; PubMed=7987400;  
 RA Reardon W., Winter R.M., Rutland P., Pulley L.J., Jones B.M.,  
 RA Malcolm S.;  
 RT "Mutations in the fibroblast growth factor receptor 2 gene cause  
 RT Crouzon syndrome";

RL Nat. Genet. 8:98-103(1994).  
 RN [9]  
 RP VARIANTS CS CYS-328 AND CYS-347 AND VARIANT JWS GLY-344.  
 RX MEDLINE=95179174; PubMed=7874170;  
 RA Jobs E.W., Li X., Scott A.F., Meyers G., Chen W., Eccles M., Mao J.,  
 RA Charnas L.R., Jackson C.E., Jaye M.;  
 RT "Jackson-Weiss and Crouzon syndromes are allelic with mutations in  
 RT fibroblast growth factor receptor 2";  
 RL Nat. Genet. 8:275-279(1994).  
 RN [10]  
 RP VARIANTS CS.  
 RX MEDLINE=95384152; PubMed=7655462;  
 RA Oldridge M., Wilkie A.O.M., Staney S.F., Poole M.D., Pulley L.J.,  
 RA Rutland P., Hockley A.D., Wake M.J.C., Goldin J.H., Winter R.M.,  
 RA Reardon W., Malcolm S.;  
 RT "Mutations in the third immunoglobulin domain of the fibroblast growth  
 RT factor receptor-2 gene in Crouzon syndrome";  
 RL Hum. Mol. Genet. 4:1077-1082(1995).  
 RN [11]  
 RP VARIANT PS ALA-321.  
 RX MEDLINE=95235551; PubMed=7719333;  
 RA Lajeunie E., Wei M.H., Bonaventure J., Munnich A., Le Merrer M.,  
 RA Renier D.;  
 RT "FGR2 mutations in Pfeiffer syndrome";  
 RL Nat. Genet. 9:108-108(1995).  
 RN [12]  
 RP VARIANTS AS TRP-252 AND ARG-253.  
 RX MEDLINE=95235562; PubMed=7719344;  
 RA Wilkie A.O.M., Staney S.F., Oldridge M., Poole M.D., Ashworth G.J.,  
 RA Hockley A.D., Hayward R.D., David D.J., Pulley L.J., Rutland P.,  
 RA Malcolm S., Winter R.M., Reardon W.;  
 RT "Apert syndrome results from localized mutations of FGR2 and is  
 RT allelic with Crouzon syndrome";  
 RL Nat. Genet. 9:165-172(1995).  
 RN [13]  
 RP VARIANTS PS PRO-341, ARG-342 AND TYR-342.  
 RX MEDLINE=95235563; PubMed=7719345;  
 RA Rutland P., Pulley L.J., Reardon W., Barstater M., Hayward R.,  
 RA Jones B., Malcolm S., Winter R.M., Oldridge M., Staney S.F.,  
 RA Poole M.D., Wilkie A.O.M.;  
 RT "Identical mutations in the FGR2 gene cause both Pfeiffer and  
 RT Crouzon syndrome phenotypes";  
 RL Nat. Genet. 9:173-176(1995).  
 RN [14]  
 RP VARIANTS CS JWS/PS.  
 RX MEDLINE=96203627; PubMed=8644708;  
 RA Meyers G.A., Day D., Goldberg R., Daentl D.L., Przylepka K.A.,  
 RA Abrams L.J., Graham J.M., Jr., Feingold M., Moeschler J.B.,  
 RA Rawnsley E., Scott A.F., Jobs E.W.;  
 RT "FGR2 exon IIIa and IIc mutations in Crouzon, Jackson-Weiss, and  
 RT Pfeiffer syndromes: evidence for missense changes, insertions, and a  
 RT deletion due to alternative RNA splicing";  
 RL Am. J. Hum. Genet. 58:491-498(1996).  
 RN [15]  
 RP VARIANTS CS CYS-105; GLY-338; CYS-351 AND ARG-384.  
 RX MEDLINE=97101656; PubMed=8946174;  
 RA Pulley L.J., Reardon W., Wilkes D., Rutland P., Jones B.M.,  
 RA Hayward R., Hall C.M., Brunton L., Chun N., Lammer E., Malcolm S.,  
 RA Winter R.M.;  
 RT "Spectrum of craniosynostosis phenotypes associated with novel  
 RT mutations at the fibroblast growth factor receptor 2 locus";  
 RL Eur. J. Hum. Genet. 4:283-291(1996).  
 RN [16]  
 RP VARIANTS CS ILE-331; ASP-336--ALA-337 DUPL AND TRP-356--THR-358 DEL.  
 RX MEDLINE=97114301; PubMed=8956050;  
 RA Steinberger D., Mulliken J.B., Mueller U.;  
 RT "Crouzon syndrome: previously unrecognized deletion, duplication, and  
 RT point mutation within FGR2 gene";  
 RL Hum. Mutat. 8:386-390(1996).  
 RN [17]  
 RP VARIANT PS CYS-290.  
 RX MEDLINE=97295073; PubMed=9150725;  
 RA Tartaglia M., Valeri S., Velardi F., di Rocco C., Battaglia P.A.;



```

11 |||
361 roblaro.....Glyargluysguliethr 370
1102 AGCGTCGCGCGCTGATCGCATCCAGCGCGCGCTCTTCA 1151
371 Alaserproaspyrleugluilealietrycylglyalphe 387
1152 CCGGAGCAGCCGCTGCTGCGCTGCGAGCGCGAGAGCGCTGCA 1201
387 uilealiacysmetvalvalthvalileucysargmetlysasntrht 404
1202 CCGCGCGCGCT 1212
404 hrlslyspro 407
seq_name: SwissProt_39:FGFR2_MOUSE
seq_documentation_block:
ID FGFR2_MOUSE STANDARD; PRT; 821 AA.
AC P21803; O61342;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE FIBROBLAST GROWTH FACTOR RECEPTOR 2 PRECURSOR (FGFR-2) (EC 2.7.1.112)
DE (KERATINOCYTE GROWTH FACTOR RECEPTOR).
GN FGFR2 OR ECT1 OR BEK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
(1)
RN RP SEQUENCE FROM N.A. (LONG FORM).
RC TISSUE=Brain, and Liver;
RX MEDLINE=92228773; PubMed=1373495;
RA Masukhant A., Dell'Era P., Moscatelli D., Kornbluth S.,
RA Hanafusa H., Basilico C.;
RT "Characterization of the murine BEK fibroblast growth factor (FGF)
RT receptor: activation by three members of the FGF family and
RT requirement for heparin."
RL Proc. Natl. Acad. Sci. U.S.A. 89:3305-3309(1992).
RN [2]
RP SEQUENCE FROM N.A. (SHORT FORM).
RX MEDLINE=91095977; PubMed=1846048;
RA Miki T., Fleming T.P., Bottaro D.P., Rubin J.S., Ron D.,
RA Aaronson S.A.;
RT "Expression cDNA cloning of the KGF receptor by creation of a
RT transforming autocrine loop."
RL Science 251:72-75(1991).
RN [3]
RP SEQUENCE OF 477-821 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89219016; PubMed=2468999;
RA Kornbluth S., Paulson K.E., Hanafusa H.;
RT "Novel tyrosine kinase identified by phosphotyrosine antibody
RT screening of cDNA libraries."
RL Mol. Cell. Biol. 8:5541-5544(1988).
CC -1- FUNCTION: RECEPTOR FOR ACIDIC AND BASIC FIBROBLAST GROWTH FACTORS.
CC POSSESSES A HIGHER AFFINITY FOR ACIDIC THAN FOR BASIC FGF'S.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: TWO FORMS OF THE PROTEIN (A LONG AND A
CC SHORT FORM) ARE PRODUCED BY ALTERNATIVE SPLICING OF THE SAME
CC GENE.
CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: M86441; AAA37286.1; -
DR EMBL: M63503; AAA39377.1; -
DR EMBL: M23362; AAA37285.1; -
DR PIR: A31378; TVMSBK.
DR PIR: A38429; A38429.
DR HSSP: P1362; 1FGI.
DR MCD: MGI:95523; Fgfr2.
DR Interpro: IPR000719; -
DR Interpro: IPR001245; -
DR Interpro: IPR003006; -
DR Pfam: PF00047; 1g; 3.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR Receptor: Glycoprotein; tyrosine-protein kinase; ATP-binding;
KW transferase; phosphorylation; transmembrane; signal;
KW Immunoglobulin domain; Alternative splicing.
FT SIGNAL 1 21
FT CHAIN 22 821
FT DOMAIN 22 377
FT TRANSMEM 378 398
FT DOMAIN 399 821
FT DOMAIN 55 114
FT DOMAIN 172 238
FT DOMAIN 271 349
FT DOMAIN 481 770
FT NP_BIND 487 495
FT BINDING 517 517
FT ACT_SITE 626 626
FT MOD_RES 657 657
FT DISULFID 62 107
FT DISULFID 179 231
FT DISULFID 278 342
FT CARBOHYD 83 83
FT CARBOHYD 123 123
FT CARBOHYD 147 147
FT CARBOHYD 228 228
FT CARBOHYD 241 241
FT CARBOHYD 265 265
FT CARBOHYD 297 297
FT CARBOHYD 318 318
FT CARBOHYD 331 331
FT VARSPPLIC 37 37
FT VARSPPLIC 38 152
FT VARSPPLIC 314 361
FT CONFLICT 169 169
FT CONFLICT 187 187
FT SEQUENCE 821 AA; 91983 MW; FCD828ADD61F4414 CRC64;

```

```

alignment_scores:
  quality: 461.00
  ratio: 1.937
Percent Similarity: 63.978
Percent Identity: 31.989

```

```

alignment_block:
US-09-598-042a-2 x FGR2_MOUSE ..
Align seg 1/1 to: FGR2_MOUSE from: 1 to: 821

```

```

112 GCGCGCATGATGCGGCTGCGATGAGTGAGGAGGAGCGCGCGCGCT 161
113 GCGCGCATGATGCGGCTGCGATGAGTGAGGAGGAGCGCGCGCGCT 161
55 Glylserleugluileuincysmetleuys.....AspAlaAla 69
162 GACCATGTGAGCAGCAGATGAGCGCGCATCCAGCGCGCTGAGCGCGCT 211
69 llsesrtrpnlrlyaspcl.....Valhlsleuglyproasnasna 84

```

```

212 TCCGCTGCTG...CCGAGGGGCTGAAGTGAACGAGTGGAGCGGAG 258
      |||||      :: |||||      ::
84  rgtHvalLeuilegIyIurIleuGlnIleYsgIyAlaIarhProarg 100
259 GATGCCGGCGTGTACGTGTGCAAGGCCAACAGCGCTTCGGAGCGCTGAG 308
      |||||      :: |||||      ::
101 AspserGIyLeuYrAlaCysThrAlaIarGthValAspserGIuTh 117
309 CGTCACTACACCCCTGCTGCTGATGACATTGAGCCAGGAGAGAGA 358
      ::      :: |||||      ::
117 rTtrIlePheMetValaInValThrAspAlaIleSerSerGIyAspAspG 134
359 GCGTGGGGCCGACAGCTCTCTGGGGGTCAAGAGACCGCCGAGCGAG 408
      :: |||||      ::
134 lAspAspThrAspSerSerGIyAspValValSerGIyAsnArgSerAsn 150
409 CAGTGGGACACAGCCGCTTCACAGCCCTCCAGATGAGCGCGCGGT 458
      || |||||      :: |||||      ::
151 GlnArGala...ProTYrThrAsnThrGIuYsMetGIuYsArgIe 166
459 GATGCCAGCGCGCTGGTAGCTCCGCGCGCTCAAGTGCCTGGCCAGCG 508
      :: |||||      :: |||||      ::
166 hIsIleYsProAlaAlaAsnThrValIlysPheArgYsProAlaGIyG 183
509 GGCACCCCTGGCGGACATCACTGATGAAGAGGACGAGCGCTTGACG 558
      |||||      :: |||||      ::
183 lYasPrOThrSerThrMetArgTrIleuYsAsnGIyIySgluPheYs 199
559 .....CGCCGAGAGCGCGCTGAGCCCGAGAGAGAGTGGACACT 599
      |||||      :: |||||      ::
200 GlnGIuHIsArgIleGIyGIyTrIySValArGAsnGlnHIsTrpSerIe 216
600 GAGCCTGAAGAACTGCGCGCGGAGACAGCGGCAATACACTGCGCGCG 649
      :: |||||      :: |||||      ::
216 uIleMetGIuSerValValProSerAspIySgluYsAnTYrThrCysIeu 233
650 TGTGCAACGGGGCGGCATCAAGCCACTACAGTGTGATGTGATC 699
      |||||      :: |||||      ::
233 AlGIuAsnGIuTrIeGIySerIleAsnHIsThrYrHIsIleuAspAlaVal 249
700 CAGCGGAGCCGCTTCACAGCCGCTGACAGCGACGACCGCGGTGAAC 749
      :: |||||      :: |||||      ::
250 GluArGserPrOHIsArgProIleuGlnAlaGIyLeuProAlaAsnAl 266
750 GACGGTGACTTCGGGGGAGACAGCTCTCCAGTGAAGTGGCGAGCG 799
      :: |||||      :: |||||      ::
266 AserThrValValGIyGIyAspValGIuPheValCysValIyTrSerA 283
800 AGCTGAACCGCGTGTATCCAGTGGCTGAAGCGGTGAGTACGCGCGCGAG 849
      |||||      :: |||||      ::
283 sPAlaGIuPrOHIsIleGIuTrIleYsHIsValGIu..... 295
850 GGCGCCCAACAATCCACCATCATGATGGCGCGCGCAGAGTTTGTG...GT 896
      :: |||||      :: |||||      ::
296 ..LysAsnGIySerIyAsnGIyProAspGIyLeuProTYrLeuIlySva 311
897 GCTGCCACAGGGTGAAGTGTGGTGGCGCGCGCAGCGCTCTCACTCAATA 946
      |||||      :: |||||      ::
311 lLeuYsAlaIaGIyVal.....AsnThrAspIySgluIleGIuIuv 326
947 AGCTGCTCATCAACCGCTGGCGCGCGCAGAGATGGCATGACATGTCG 996
      |||||      :: |||||      ::
326 AlleuTYrTrIleArGAsnValThrPheGIuAspAlaGIyGIuYrThrYs 342
997 CTTGGCGCAACACCATGGGTACACTTTCGACAGCGCTTCCTCACTGAGT 1046
      |||||      :: |||||      ::
343 LeuAlaGIuAsnSerIleGIyIleSerPheHIsSerAlaTrIlePheThVa 359
1047 GCTGCCAACAACCAACGCGGAGCGGACGCTGTGGCTCTGCTGCTGCG 1096
      |||||      :: |||||      ::
359 lIleuProAlaPro.....ValArGIuYsGIuI 369
1097 CCACrAGAGCTGGCGCGGTGTCATCGGATCCACGCGCGCTGTC 1146

```

```

369 lEthrAlaSerProAspTrIyLeuGIuIleAlaIleTYrCysIleGIyAl 385
1147 TTCATCTGGGACACCTGCTCTGTGGCTTTCGACAGCCCGCAGAGAGCC 1196
      |||||      :: |||||      ::
386 PheIleuIleAlaCysMetValValThrValIlePheCysArgMetIySth 402
1197 GTCACACCCCGCGCT 1212
      |||||      ::
402 rThrTrIyIySPro 407
seq_name: SwissProt_39:FGFR4_HUMAN
seq_documentation_block:
ID FGFR4_HUMAN STANDARD; PRT; 802 AA.
AC P22455;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE FIBROBLAST GROWTH FACTOR RECEPTOR 4 PRECURSOR (FGFR-4) (EC 2.7.1.112).
GN FGFR4 OR JTK2 OR TKF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91224085; Pubmed=1709094;
RA Partanen J.M., Maekelae T.P., Eerola E., Korhonen J., Hirvonen H.,
RT Claesson-Welsh L., Alltalo R.;
RT "FGFR-4, a novel acidic fibroblast growth factor receptor with a
RT distinct expression pattern.";
RL EMBO J. 10:1347-1354(1991).
RN [2]
RP SEQUENCE OF 609-676 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=91062389; Pubmed=2247464;
RA Partanen J., Maekelae T.P., Alltalo R., Leheslahti H., Alltalo K.;
RT "Putative tyrosine kinases expressed in K-562 human leukemia cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8913-8917(1990).
CC -1- FUNCTION: RECEPTOR FOR ACIDIC FIBROBLAST GROWTH FACTOR. DOES NOT
CC -1- BIND TO BASIC FIBROBLAST GROWTH FACTOR. BINDS FGF19.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X57205; CAA40490.1; -;
DR EMBL; M59373; AAA63208.1; -;
DR PIR; S15345; TVH0F4.
DR HSSP; P11362; IFGI.
DR MIM; 134935; -;
DR InterPro; IPR000719; -;
DR InterPro; IPR001245; -;
DR InterPro; IPR003006; -;
DR Pfam; PF00047; 1g; 3.
DR PRINTS; PR00069; PKINASE.1.
DR PRINTS; PR00109; TYRKINASE.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP_1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM_1.
DR Receptor; glycoprotein; tyrosine-protein kinase; ATP-binding;
KW Transferase; Phosphorylation; Transmembrane; Immunoglobulin domain;
KW Signal.
FT SIGNAL 1 24 POTENTIAL.

```

```

FT CHAIN 25 802 FIBROBLAST GROWTH FACTOR RECEPTOR 4.
FT DOMAIN 25 369 EXTRACELLULAR (POTENTIAL).
FT TRAMSM 370 390 POTENTIAL.
FT DOMAIN 391 802 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 50 108 IG-LIKE DOMAIN.
FT DOMAIN 165 231 IG-LIKE DOMAIN.
FT DOMAIN 264 340 IG-LIKE DOMAIN.
FT DOMAIN 467 755 PROTEIN KINASE.
FT NP_BIND 473 481 ATP (BY SIMILARITY).
FT BINDING 503 503 ATP (BY SIMILARITY).
FT ACT_SITE 612 612 BY SIMILARITY.
FT MOD_RES 643 643 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT DISULFID 57 101 POTENTIAL.
FT DISULFID 172 224 POTENTIAL.
FT DISULFID 271 333 POTENTIAL.
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 258 258 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 311 311 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 322 322 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 802 AA: 87938 MW: B2C654A31BB89F CRC64;

```

```

alignment_scores:
    Quality: 459.50      Length: 409
    Ratio: 1.947        Gaps: 15
    Percent Similarity: 57.702    Percent Identity: 31.785

```

alignment\_block:

US-09-598-042a-2 x FGR4\_HUMAN

Align seg 1/1 to: FGR4\_HUMAN from: 1 to: 802

```

109 CTGGCGCCGACATAGCGGTGACAGTGGCCAGTGAGGGGAGCCGCCGCC 158
    |||||: : : : ||||| |||
49 LeuGlyGlnProValArgLeuGlyCys.....GlyArgAlaGluArg 62
159 GCTGACCATGTGACACCAAGATGAGCCGACACCAACAGCGGTGAGCC 208
    ||| |||||: : : : |||
62 GgLyGlyHisTrpTyrLysGluGlySer..... 71
209 GCTTCGCGGTGCTGCGCAGGGG.....CTG 234
    |||||: : : : |||
72 ....ArgLeuAlaProAlaGlyArgValArgIlyTrpArgIlyArgLeu 86
235 AAGGTGAAGAGTGAGGCGGAGATGCGGGGTGACGGGTGACGAGCC 284
    : : : : : ||||| ||||| ||||| |||
87 GluLeuLeuAspSerPheLeuProGluAspAlaGlyArgTyrLeuGlySer 103
285 CACCAACGCGCTTCGCGCAGCTGAGCTC.....AACTACACCTCGTCG 328
    | |||||: : : : ||| |||||: : : :
103 Arg.....GlySerMetIleValLeuGlnAsnLeuThrLeuIleT 117
329 TGTGTGATGACATTAGCCAGGAGAGAGAGCTGGGGCCCGACAGCTCC 378
    : : |||||: : : : : ||| |||
117 hrgIyAspSerLeuThrSerSerAsnAspAspGluAspProLysSer... 132
379 TCTGGGGGTACAGAGAGCCCGCC.....AGCCAGAGTGGGACG 419
133 .....HisArgAspProSerAsnArgHisSerTyrProGlnGlnAl 146
420 ACCGCGCTTCACACAGCCCTCCAGATGAGCGCGGGGTGTCGACGCGC 469
    ||| : : |||||: : : : ||| |||
146 aProTyrTrpThrHisProGlnArgMetGlyLysLeuHisAlaValP 163
470 CCGTGGGTAGCTCCGTCGGCTCAAGTCGCGGCGACGCGGACCTCGG 519
    |||||: : : : ||| |||||: : : : |||
163 roAlaGlyAsnThrValLysPheArgCysProAlaAlaGlyAsnProThr 179
520 CCGGACATCAGTGGATGAGAGGACGACGCGCTTG.....ACGCG 560
    ||| ||| |||||: : : : ||| |||||: : : :
180 ProThrIleArgTrpLeuLysAspGlyGlnAlaPheHisGlyGluAsnAr 196

```

```

561 CCCAGAGCCGCTGAGCCAGAGAGAGAGTGAACACTGAGCTGAGAGA 610
    | : : : : ||| : : : |||||: : : :
196 gIleGlyGlyIleArgLeuArgHisGlnHisTrpSerLeuValMetGlu 213
611 ACCGTGGGCGGAGAGACAGCGGCAATACCTGCCGCGGTGTCGAAAGC 660
    : : : : : ||||| ||| |||||: : : : |||
213 erValValProSerAspArgGlyThrTyrThrCysLeuValGlnAsnAla 229
661 GGGGGGCGGATCAGCCGACCTCAAGAGGTGATGTGATCAGCGAGCCG 710
230 ValGlySerIleLeuTyrAsnTyrLeuLeuAspValLeuGlnArgSer 246
711 TTCAGAGCCGCTGTCTACAGCAGCAGCCCGCTGAAACAGAGGTGACT 760
    : : : : : ||||| ||| |||||: : : : |||
246 ohnAspProIleLeuGlnAlaGlyLeuProAlaAsnThrAlaVal 263
761 TGGGGGAGACACGCTCTTCCAGTGAAGTGGCAGCAGCTGAAAGCG 810
    |||||: : : : : ||||| |||||: : : : |||
263 aGlySerAspValGlnLeuLeuGlyValTyrSerAspAlaGlnPro 279
811 GTGATCAGTGGCTGAGCGCGTG.....GAGTACGCGCG 845
    |||||: : : : : ||||| |||||: : : : |||
280 HisIleGlnTrpLeuLysHisIleValIleAsnGlySerPheGlyAl 296
846 CGAGGCGCCGACCACTCCACATGATGATGGCGCGCGCAAGATTGTG 895
    | : : : : |||||
296 a.....ValGlyPheProTyrValGlnVal 304
896 TGTGCGCCAGGGGTGACGCTGTGTCGCGCGCGGCTCTACCTCAT 945
    ||||| |||||: : : : ||||| |||||: : : : |||
304 alLeuLysThrAlaAspIle.....AsnSerSerGluValGlu 316
946 AAGCTCTCATCACCCTGCGCGCGGAGAGATGCGGCGCATGACATCTG 995
    ||| : : : : : ||||| ||||| |||
317 ValLeuTyrLeuArgAsnValSerAlaGluAspAlaGlyIlyTyrThr 333
996 CCTGGCGCCACACCATATGGGTGACAGCTTCCGACGCGCTTCTACCG 1045
    |||||: : : : : ||||| |||||: : : : |||
333 sleuAlaGlyAsnSerIleGlyLeuSerTyrGlnSerAlaTrpLeuThr 350
1046 TGCTGCAGACCCCAAAACGCGGCGGACGACCTGTGCTCTCTCTCTG 1095
    |||||: : : : : ||| |||||: : : : |||
350 alLeuProGlnGluAspProThrTrpThrAlaAlaAlaProGlnAlaArg 366
1096 GCCATAGCTGCGCGGTGCGCGGTGATCGGATCGGATCCAGCGGCTGT 1145
    |||||: : : : : ||||| |||||: : : : |||
367 TyrThrAsp.....IleIleLeuTyrAlaSerIlySerIle 378
1146 CTTCATCTGCGGACACCTGCTCTG...TGCTTTCGACAGCGCCAGAGA 1192
    : : : : : ||||| ||| |||||: : : : |||
378 uAlaLeuAlaValLeuLeuLeuLeuAlaGlyLeuTyrArgGlyGln.... 393
1193 AGCGGTGACACCCCGCGCTGCGCTCCCTGCGCTGGGCGAC.....CGC 1236
394 .....AlaLeuHisGlyArgHisProArg 401
1237 CCGCGGGGAGCGCGCGGACGCGAGC 1263
    |||||: : : : : |||
402 ProProAlaThrValGlnLysLeuSer 410
seq_name: SwissProt_39:CEK2_CHICK
seq_documentation_block:
ID CEK2_CHICK STANDARD; PRT; 806 AA.
AC P18460;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE TYROSINE KINASE RECEPTOR CEK2 PRECURSOR (EC 2.7.1.112).
GN CEK2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

```



OC Gallus.  
 OX NCBI\_TaxID=9031.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90332672; PubMed=2165604;  
 RA Pasquale E.B.;  
 RT "A distinctive family of embryonic protein-tyrosine kinase  
 receptors."  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:5812-5816(1990).  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR  
 CC FAMILY.  
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M35195; AAA48664.1; -  
 DR PIR: A35963; A35963.  
 DR HSSP: P11362; 1FGI.  
 DR InterPro: IPR000719; -  
 DR InterPro: IPR001245; -  
 DR InterPro: IPR003006; -  
 DR Pfam: PF00047; 1g; 3.  
 DR Pfam: PF00069; pkinase; 1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR Receptor: Glycoprotein; Tyrosine-protein kinase; ATP-binding;  
 KM Transferrase; Phosphorylation; Transmembrane; Immunoglobulin domain;  
 KM Signal.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 806 TYROSINE KINASE RECEPTOR CER2.  
 FT DOMAIN 20 364 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 365 389 POTENTIAL.  
 FT DOMAIN 390 806 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 54 114 IG-LIKE DOMAIN.  
 FT DOMAIN 163 229 IG-LIKE DOMAIN.  
 FT DOMAIN 262 340 IG-LIKE DOMAIN.  
 FT DOMAIN 131 141 ASP/GLU-RICH (HIGHLY ACIDIC).  
 FT DOMAIN 466 755 PROTEIN KINASE.  
 FT DOMAIN 472 480 ATP (BY SIMILARITY).  
 FT BINDING 502 502 ATP (BY SIMILARITY).  
 FT ACT\_SITE 611 611 BY SIMILARITY.  
 FT MOD\_RES 642 642 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT DISULFID 61 107 POTENTIAL.  
 FT DISULFID 170 222 POTENTIAL.  
 FT DISULFID 269 333 POTENTIAL.  
 FT CARBOHYD 83 96 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 96 96 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 118 118 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 219 219 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 256 256 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 288 288 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 309 309 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 322 322 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 806 AA; 89730 MW; B3BB3C6D5F231AB6 CRC64;

## alignment\_scores:

Quality: 456.50 Length: 440  
 Ratio: 1.723 Gaps: 13  
 Percent Similarity: 60.227 Percent Identity: 30.227

## alignment\_block:

US-09-598-042A-2 x CER2\_CHICK ..

Align seg 1/1 to: CER2\_CHICK from: 1 to: 806

```

112 GCGCCGACTATGCGGCTGCAAGTCCCAAGTGAAGGAGACCCGCCGCT 161
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
54  GLYSPNTHrIleGluLeuSerCysAsnThrGlnSerSerValSerVa 70
162 GACCATGTGGACCAAGATGGCCGACCATCCACAGCGGCTGACCCCT 211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
70  I...PheThrPheLysAspGlyIleGlyIleAlaProSer...AsnArgT 85
212 TCCGCGTCTGCGCGGCTGCAAGTGAAGCAGGTGAGCGGAGGAT 261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
85  hrhIleGlyIleGluLeuLeuLeuIleIleAsnValSerIleAsp 101
262 GCGCGGTGTGATGTCAGTGCAGGCGCACCAAGCGCTTGCGAGCTGAG 311
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
102 SerGlyLeuIleYserCysLysProArgHisSerAsnGluValLeuGly.. 117
312 CACATACACCCCTGCTGCTGCTGATGATGATGATGATGATGATGATG 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
118  AsnPheThrValAlaArgValThrAsp..... 125
362 TGGGCGCGGACGCTCTCTGCGGCTGCAAGAGACCCGCCAGCAGAG 411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
126  SerProSerSerIleYspAspGluAspAspAspAspGluSerGluAsp 141
412 TGGGACGACCGCGCTTCAACAGCCCTCCAGATGAGCGCGCGGTAT 461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
142 ThrGlyValProPheThrPheArgProAspLysMetGluLysLeuLeu 158
462 CGCAGCGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 511
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
158 ValAlaValProAlaAlaAsnThrValAlaArgPheArgPheProAlaGly 175
512 ACCCTGCGCGCGCATCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAG 558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
175 snProThrProThrIleThrIleThrIleThrIleThrIleThrIleThr 191
559 .....GCGCCAGAGCGGCTGAGCCAGAGAGAGAGAGAGAGAGAGAG 602
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
192 GluHisArgIleGlyIleLysLeuArgHisGlnGlnGlnGlnGlnGln 208
603 CGTGAAGAACCTGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 652
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
208 IMetIleSerValValProSerValArgLysAsnThrValValValG 225
653 CGAACCGCGCGCGCGCATCAAGCGGCGGCGGCGGCGGCGGCGGCG 702
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
225 IuAsnLysIleIleLysAsnIleArgHisIleThrIleGluAspValLeuGlu 241
703 CGGACCGCGTCCAGACCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCT 752
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
242 ArgSerProHisArgProIleLeuGlnAlaGlyLeuProAlaAsnGlnTh 258
753 GGTGACATTCGGGGGACACAGCTCTTCCAGTGAAGTGGCGGACGAG 802
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
258 IValValValGlySerAsnValGluPheHisCysLysValYserAspA 275
803 TGAACCGCGGTATCCAGTGGCGGAGCGGTGAG..... 837
275 IagInProHisIleGlnIlePheLysHisValGluValAsnGlySerLys 291
838 TACGCGCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 887
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
292 TyrGlyProAspGlyThrProThrValThr..... 301
888 GTTGTGCTGCTGCGGCGGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCT 937
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
302 .....ValLeuLysThrAlaGlyVal.....AsnThrThrAspLysG 314
938 ACCTCAATAGCTGCTCATACCGCGGCGGCGGCGGCGGCGGCGGCGG 987
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
314 IuLeuGluIleLeuThrIleuArgAsnValThrPheGluAspAlaGlyGlu 330

```

```

988 TACATGCTGCTGGCGCAACACCATGGCTACGCTTCGACGCCCTT 1037
||| |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
331 TyThcysenualaglynsenSerIleGlyPheSerHisHisSerIaltr 347
||| |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1038 CCTCACCCTGCTGCGACACCAACCGCCAGGCCACCTGTGGCTCTCT 1087
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
347 PleuThrValLeuProAlaGluIleuMetGluMetAspSerGlyS 364
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1088 CGTTCCTGGCCACTAGCTGCGCCCTGCTCATCGCATCCAGCC 1137
||| |||||:|||||:|||||:|||||:|||||:|||||:|||||:
364 eValTyAlaGlyIleLeuSerTyGlyThrGlyLeu..... 376
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1138 GGGGCTGCTCATGCTGGCGACCTGCTGCTGGCTGCTGGCGCCA 1187
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
377 ...ValIleuHeIleuValIleuValIleValIleIleGlySerG...Me 391
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1188 GAAGAACCCG.....TCACCCCGCCGCGCTGCC 1216
||| |||||:|||||:|||||:|||||:|||||:|||||:|||||:
391 TlysmetProAsnIlyslsAlaMetAsnThrThrIleValGlnIlyValS 408
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1217 CTCCTGCTGCTGGCGACCCCGCGCGCGCCGCGACCCAGC... 1263
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
408 eRlyspHeProLeuLysArgGlnGlnValSerLeuGlnSerAsnSer 424
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1264 .....GGAGACAAAGACCTTCCTGCTGCGCGCGCTGCGCTGCC 1307
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
425 MetAsnSerAsnThrProLeuValArgIleThrArgLeuSerSerAs 441
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1308 TGGTGTGGGGCTGTGTAGAGAGCATGGGTCTCGGCGACCCCGACACT 1357
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
441 pGlyProMetLeuAlaAsnValSerGluLeuGlnIleuProAspPro 457
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1358 TACTGGGCCCAAGGCCCGCAG 1375
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
458 LysTrpGluLeuAlaArg 463
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
seq_name: swissprot_39:FGR4_MOUSE
seq_documentation_block:
ID FGR4_MOUSE STANDARD: PRT; 808 AA.
AC 003142;
DT 01-OCT-1994 (rel. 30, Created)
DT 01-OCT-1996 (rel. 34, Last sequence update)
DT 15-JUL-1999 (rel. 38, Last annotation update)
DE FIBROBLAST GROWTH FACTOR RECEPTOR 4 PRECURSOR (FGFR-4) (EC 2.7.1.112)
DE (PROTEIN-TYROSINE KINASE RECEPTOR MPK-11).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RN SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Petal cerebellum;
RX MEDLINE=92146274; PubMed=1723680;
RA Stark K.L., McMahon J., McMahon A.P.;
RA "FGFR-4, a new member of the fibroblast growth factor receptor
RT family, expressed in the definitive endoderm and skeletal muscle
RT lineages of the mouse."
RL Development 113:641-651(1991).
RN (2)
RN SEQUENCE OF 620-676 FROM N.A.
RX STRAIN=C57BL; TISSUE=Embryonic brain;
RX MEDLINE=93096484; PubMed=1281307;
RA Gilaudi-Hedensstreit P., Nleto M.A., Frain M., Mattei M.-G.,
RA Chester A., Wilkinson D.G., Charney P.;
RA "An Eph-related receptor protein tyrosine kinase gene segmentally
RT expressed in the developing mouse hindbrain."
RL Oncogene 7:2499-2506(1992).
-1- FUNCTION: PUTATIVE RECEPTOR FOR BASIC FIBROBLAST GROWTH FACTOR.
CC MAY BE INVOLVED IN THE DEVELOPMENT OF SKELETAL MUSCLE CELL
CC LINES.

```

```

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE DEVELOPING GUT ENDODERM,
CC IN MYOTOMALLY DERIVED SKELETAL MUSCLE, THE ADRENAL CORTEX,
CC KIDNEY AND CONDENSING CARTILAGE. PRESENT IN ADULT LIVER.
CC LUNG AND KIDNEY.
CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X59927; CAA42551.1; -.
DR EMBL; X57236; CAA40512.1; -.
DR HSSP; P11362; IFGI.
DR MGD; MGI:95525; Fgfr4.
DR InterPro; IPR000719; -.
DR InterPro; IPR001245; -.
DR InterPro; IPR003006; -.
DR Pfam; PF00047; Ig; 3.
DR Pfam; PF00069; PKinase; 2.
DR PRINTS; PR00109; TYRKINASE.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM_1.
DR KMW; KMW000001; Tyrosine-protein kinase; ATP-binding;
KW Transferase; Glycophorin; Tyrosine-protein kinase; Immunoglobulin domain;
KW Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 808 FIBROBLAST GROWTH FACTOR RECEPTOR 4.
FT DOMAIN 19 366 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 367 387 POTENTIAL.
FT DOMAIN 388 808 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 53 98 IG-LIKE DOMAIN.
FT DOMAIN 169 221 IG-LIKE DOMAIN.
FT DOMAIN 268 330 IG-LIKE DOMAIN.
FT DOMAIN 464 761 PROTEIN KINASE.
FT NP_BIND 470 478 ATP (BY SIMILARITY).
FT BINDING 509 509 ATP (BY SIMILARITY).
FT ACT_SITE 618 618 BY SIMILARITY.
FT MOD_RES 649 649 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT DISULFID 53 98 POTENTIAL.
FT DISULFID 169 221 POTENTIAL.
FT DISULFID 268 330 POTENTIAL.
FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 319 319 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 808 AA; 89775 MW; E3F30BE5B34E8E6 CMC64;

alignment_scores:
Quality: 442.00 Length: 455
Ratio: 1.720 Gaps: 23
Percent Similarity: 56.484 Percent Identity: 32.527

alignment_block:
US-09-598-042A-2 x FGR4_MOUSE ..
Align seg 1/1 to: FGR4_MOUSE from: 1 to: 808
109 CTGGCGCCGACTATGGCGCTGACGTGACGATGGAGGGAGCCGCCGCC 158
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
46 LeuGlyGlnProValArgLeuGlyCys.....GlyArgThrGluar 59
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
159 GCGACCATGTGACCAAGATGGCGCGCACCATCCACAGC..... 198
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

59  gGlyArgHisTrpTyrLysGluGlySerArgLeuAlaSerAlaGlyArgV 76
199  .... GCGTGAAGCCGCTTCGCGCTGTCGCCGACGAGCGCTGAAGTGAAG 243
76  AlaArgLysTrp... ArgGlyArg..... LeuGluLeuAla 86
244  CAGGTGAGCGGAGGAGTGGCGCTGACGTGACGTGACGACGACGACGAGG 293
87  SerPheLeuProGluAspAlaGlyArgTyrLeuLysLeuAlaArg.... 101
294  CTTCGCGACGCTGACGCTCACTACACCTGCTGCTGCTGATGAC.... 339
102  ... GlySerMetThrValValHisAsnLeuThrLeuLeuMetAspAsp 117
340  .... ATTAGCCGAGGGAAGAGAGCGCTGGGCGCGACAGCTCC 378
117  erLeuThrSerLeuSerAsnAspGluAspProLysThrLeuSerSerSer 133
379  TCTGGGGTCAAGAGAGACCGCGCCGACGACGAGTGGGACGACCGCGCT 428
134  SerSerGlyHisValTyrPro.... GlnGln.... AlaProTyrTr 146
429  CACACAGCGCTCCAGATGAGCGCGCGGTGATCGACGCGCGCGGTGTA 478
146  PThrHisProGlnArgMetGluLysLysLeuHisAlaValProAlaGlyA 163
479  GCTCCGTCGCGCTCAAGTGCCTGGCGACGCGGACCGCTGCGCGACATC 528
163  snThrValLysPheArgCysProAlaCysArgAsnProMetProThrIle 179
529  ACGTGGATGAGAGAGACGACGCGCTTG..... ACGCGCGCGACGAGC 569
180  HisTrpLeuLysAspGlyGlnAlaPheHisGlyLysAsnArgLysGly 196
570  CGCTGAGCGCCGAGAGAAAGTGCACACTGACCTGAAGAACTGCGCGC 619
196  YLLeuArgLeuArgHisGlnHisTrpSerLeuValMetGluSerValValP 213
620  CGAGAGAGCGCGCAATATACACTCCGCGCTGTCGACCGCGCGCGCGCC 669
213  roseAspArgGlyThrTyrThrCysLeuValGluAsnSerLeuLysSer 229
670  ATCAACGCGCACTCAAGTGATGTATCCAGCGGACCGCGCTCCAAACC 719
230  LLeuArgTyrSerTrpLeuLeuAspAlaLeuGluArgSerProHisArgPr 246
720  CGTGTCTACAGACGACCGCGCTGACGACGAGCGGTGACCTTCGGGGGA 769
246  oLLeuGlnAlaGlyLeuProAlaAsnThrThrAlaValAlaGlySerA 263
770  CCACGCTCTTCCACTGCAAGGTGCGCAGCGACGACGCTGATGTCAG 819
263  spValGluLeuLeuLysValTyrSerAspAlaGlnProHisLysGln 279
820  TGGCTGAAGCGCGGTGAGTACGCGCGCGACGCGCGCACTCCACCAT 869
280  TrpLeuLysHisValValIle..... AsnGlySerSerPh 291
870  CGATGTGCGCGCGAGAGTTGTG... GTGCTGCCACGCGGTGACGCTG 916
291  eGlyAlaAspLysPheProTyrValGlnValLeuLysThrThrAspIle. 307
917  GGTGCGCGCGCGCGCTCTACTCAATATGACGCTGATCACCCTGCTCC 966
308  .... AsnLeuSerGluValGlnValLeuLysThrArgAsnVal 320
967  CGCCAGAGCAATGCGGGCATGTACTGCTTGGCGCAACACCATGGG 1016
321  SerAlaGluAspAlaGlyGluTyrThrCysLeuAlaGlyAsnSerIleG 337
1017  CTACAGCTTCGCGACGCGCTTCCTACACGCTGCTGCACGACCAAAACGC 1066
337  YLeuSerTyrGlnSerAlaTrpLeuThrValLeuProGlu..... 350

```

```

1067  CAGGCGCACTGTGGCTTCCTGCTGCGCCACTAGCCTGCGCTG... 1113
351  .... GlnAspLeuThrTrpThr 356
1114  .... CCGGNGCATCGGATCCGACG 1136
357  ThrAlaThrProGlnAlaArgTyrThrAspIleIleLeuTyrValSerG 373
1137  CGGCGCTGTCTTCATCTGCGGACCGCTGCTGCTGCG... 1174
373  YSerLeuValLeuLeuValLeuLeuLeuLeuAlaGlyValTyrHisArg 389
1175  ...TTTGCAGGCGCCAGAGAAGACCGCTGCACCC... 1204
390  GlnValIleArgGlyHisTyrSerArgGlnProValThrIleGlnLysLe 406
1205  .CGGCGCTGCGCGCTCCCTGCTGCGGACGCGCGCGGAGCGCGCG 1253
406  userArgPheProLeu..... AlaArgGln..... 414
1254  CGACCGCAGCGGAGACAAGACCTTCCTGTTGGCGCGCTCAGCGCTG 1303
415  .PheSerLeuGluSerArgSerArgLys..... SerSerLeu 427
1304  GCCCTGTGTGGGCTGTGTGAGAGACATGGGTCCGCGACCGCGCAG 1353
428  SerLeuValAlaArgGly... ValArgLeuSerSer... Proprol. 442
1354  CACTTACTGGG 1364
442  euLeuThrGly 445

seq_name: SwissProt_39:FGRL_CHICK
seq_documentation_block:
ID FGRL_CHICK STANDARD; PRT; 819 AA.
AC P21804;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-JUN-1999 (Rel. 38, Last annotation update)
DE BASIC FIBROBLAST GROWTH FACTOR RECEPTOR 1 PRECURSOR (EC 2.7.1.112).
GN CER1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89315814; Pubmed=2473471;
RA Pasquale E.B., Singer S.J.;
RT "Identification of a developmentally regulated protein-tyrosine
RT kinase by using anti-phosphotyrosine antibodies to screen a cDNA
RT expression library.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:5449-5453(1989).
RN [2]
RP REVISIONS.
RA Pasquale E.B.;
RL Submitted (MAY-1989) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89298406; Pubmed=2544996;
RA Lee P.L., Johnson D.E., Cousens L.S., Fried V.A., Williams L.T.;
RT "Purification and complementary DNA cloning of a receptor for basic
RT fibroblast growth factor.";
RL Science 245:57-60(1989).
CC -!- FUNCTION: RECEPTOR FOR BASIC FIBROBLAST GROWTH FACTOR.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
CC FAMILY.
CC -!- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE DOMAINS.

```

[illegible]

```

1126 GGCATCCAGCGCGCGCTGCTCATCTGCGGACCCGCTGCTGCT 1175
1176 TTGCGAGCGCCGAGAGAGCCGTCGACCCCGCGCTC... 1212
392 Lllellyrlysmetylserserthrlhlyslsyrhsrphsnaenserg 409
1213 .....GCCCCCTCCCTGCGCGGACCCGCGCGC 1242
409 lnleuAlaValAlhlsylsleuAlatylsSerlleProleuArgrglnVal 425
1243 GGGAGCGCGCCGACCGGACGCGAGACAAGAC..... 1275
426 ThrValSerAlaAspserSerSerSerMetAsnSerGlyValMetLeuVa 442
1276 .CTTCCCTGCTGTCGCGCCCTCAGCGCTGCTGTCGTCGTCG 1324
442 lArgProSerArgleuSerSerSerGlyThrProMetLeuAlaGlyValS 459
1325 AGGAGCATGGCTCCGCGGACCCCGCAG 1353
459 erglutyrgluleuProgluAspProarg 468

seq_name: SwissProt_39:FGRL_MOUSE

seq_documentation_block:
ID FGRL_MOUSE STANDARD; PRT; 822 AA.
AC P16092; Q01736; Q61562;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1991 (Rel. 18, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE BASIC FIBROBLAST GROWTH FACTOR RECEPTOR 1 PRECURSOR (BFGF-R)
   (EC 2.7.1.112) (MFR).
GN FGFR1 OR FLT3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID:10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RA MEDLINE-90160373; PubMed-1689490;
RA Reid H.H., Wilks A.F., Bernard O.;
RT "Two forms of the basic fibroblast growth factor receptor-like mRNA
RT are expressed in the developing mouse brain.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:1596-1600(1990).

[2]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-Brain;
RA MEDLINE-90265603; PubMed-2161096;
RA Saffran A., Avivi A., Orr-Urtreger A., Neufeld G., Lonal P.,
RA Givol D., Yarden Y.;
RT "The murine flg gene encodes a receptor for fibroblast growth
RT factor.";
RL Oncogene 5:635-643(1990).

[3]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC MEDLINE-91207411; PubMed-1708247;
RA Kounhara H., Kasayama S., Saito H., Matsumoto K., Sato B.;
RT "Expression cDNA cloning of fibroblast growth factor (FGF) receptor
RT in mouse breast cancer cells: a variant form in FGF-responsive
RT transformed cells.";
RL Biochem. Biophys. Res. Commun. 176:31-37(1991).

[4]
RP SEQUENCE FROM N.A. (SHORT FORM).
RC MEDLINE-90272715; PubMed-2161540;
RA Mansukhani A., Moscattelli D., Talarico D., Levyska V., Basilico C.;
RT "A murine fibroblast growth factor (FGF) receptor expressed in CHO
RT cells is activated by basic FGF and Kaposi FGF.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:4378-4382(1990).

[5]

```

```

RP SEQUENCE OF 1-15 FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE-95100926; PubMed-7802632;
RA Harada T., Saito H., Kounhara H., Kurebayashi S., Kasayama S.,
RA Terakawa N., Kishimoto T., Sato B.;
RT "Murine fibroblast growth factor receptor 1 gene generates multiple
RT messenger RNAs containing two open reading frames via alternative
RT splicing.";
RL Biochem. Biophys. Res. Commun. 205:1057-1063(1994).
CC -1- FUNCTION: RECEPTOR FOR BASIC FIBROBLAST GROWTH FACTOR. A SHORTER
CC FORM OF THE RECEPTOR COULD BE A RECEPTOR FOR ACIDIC FGF (aFGF).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: A LONG FORM (SHOWN HERE), A
CC VARIANT FORM AND A SHORT FORM ARE PRODUCED BY ALTERNATIVE
CC SPLICING.
CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M28998; AAA37290.1; -
DR EMBL; X51893; CAA36175.1; -
DR EMBL; M65053; AAA37620.1; -
DR EMBL; M33760; AAA37622.1; -
DR EMBL; S74765; AAB32845.1; ALT_SEQ.
DR PIR; A34849; TVMSFG.
DR HSSP; P11362; IFGI.
DR MGDI; MGI:95522; Fgfr1.
DR InterPro; IPR000719; -
DR InterPro; IPR001245; -
DR InterPro; IPR003006; -
DR Pfam; PF00047; Ig; 3.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
DR KMW; Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;
DR KMW; Transferrase; Phosphorylation; Transmembrane; Signal;
KW Immunoglobulin domain; Alternative splicing.
FT SIGNAL 1 21
FT CHAIN 22 822
FT DOMAIN 22 376
FT TRANSMEM 377 397
FT DOMAIN 398 822
FT DOMAIN 48 108
FT DOMAIN 171 237
FT DOMAIN 270 348
FT DOMAIN 478 767
FT NP_BIND 484 492
FT BINDING 514 514
FT ACT_SITE 623 623
FT MOD_RES 654 654
FT DISULFID 35 101
FT DISULFID 117 227
FT DISULFID 178 230
FT CARBOHYD 117 341
FT CARBOHYD 227 77
FT CARBOHYD 240 227
FT CARBOHYD 264 264
FT CARBOHYD 296 296
FT CARBOHYD 317 317
FT CARBOHYD 330 330
FT VARSPIC 30 30
FT VARSPIC 31 119

```



[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE-93201992; PubMed-1284237;  
RA Friesel R., Brown S.A.N.;  
RT "Spatially restricted expression of fibroblast growth factor  
receptor-2 during xenopus development.";  
RL Development 116:1051-1058(1992).  
CC -1- FUNCTION: RECEPTOR FOR ACIDIC AND BASIC FIBROBLAST GROWTH  
FACTORS.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE ANTERIOR NEURAL PLATE  
IN EARLY NEURULA STAGE EMBRYOS. LATER IN DEVELOPMENT, THE  
CC PROTEIN IS ALSO EXPRESSED IN THE EYE ANLAGEN, MIDBRAIN-HINDBRAIN  
CC BOUNDARY AND OTIC VESICLE.  
CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR  
FAMILY.  
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X65943; CAA46758.1; -.  
CC PIR: S25060; S25060.  
CC DR HSSP: P11362; 1FGI.  
CC DR InterPro: IPR000719; -.  
CC DR InterPro: IPR001245; -.  
CC DR InterPro: IPR003006; -.  
CC DR Pfam: PF00047; 1g; 3.  
CC DR Pfam: PF00069; tyrase; 1.  
CC DR PRINTS: PR00109; PKKINASE.  
CC DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
CC DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
CC DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
CC DR PROSITE: PS00111; PROTEIN\_KINASE\_TYR; 1.  
CC KW Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;  
CC Transferrase; Phosphorylation; Transmembrane; Immunoglobulin domain;  
CC Signal.  
CC KW  
FT SIGNAL 1 14  
FT CHAIN 15 813  
FT DOMAIN 18 367  
FT TRANSMEM 368 388  
FT DOMAIN 389 813  
FT DOMAIN 471 760  
FT DOMAIN 51 106  
FT DOMAIN 163 228  
FT DOMAIN 261 339  
FT NP\_BIND 477 485  
FT BINDING 507 507  
FT ACT\_SITE 616 616  
FT MOD\_RES 647 647  
FT CARBOHYD 79 79  
FT CARBOHYD 115 115  
FT CARBOHYD 231 231  
FT CARBOHYD 255 255  
FT CARBOHYD 287 287  
FT CARBOHYD 308 308  
FT CARBOHYD 321 321  
SQ SEQUENCE 813 AA; 91340 MW; 815436569892A565 CRC64;

## alignment\_scores:

Quality: 425.50 Length: 409  
Ratio: 1.780 Gaps: 14  
Percent Similarity: 58.435 Percent Identity: 29.584

## alignment\_block:

US-09-598-042A-2 x FGR2\_XENLA ..  
Align seg 1/1 to: FGR2\_XENLA from: 1 to: 813

112 GCGCCGACATATCGGCTGACATGCGCCAGTGGAGGCGGACCCGCGCGCT 161  
||||| :  
51 GlytupProleuAspLeuArgCysProleuAlaAspGly.....Prole 65  
162 GACCATGTGGACCAAGATGGCGCCAGCATCCACAGCGCGCTGGAGCCGCT 211  
:  
65 uValThrTrpThrLysAspGlyAlaLysLeuGluValasn...AsnArgT 81  
212 TCGCGTGTGCGCGGAGGCTGAAGGTGAAGCAAGTGGAGCGAGAGAGT 261  
:  
81 hrLeuIleValArgThrLysLeuGlnIleLysGluSerThrThrArgasp 97  
262 GCGCGCTGTACGTGTGCAAGGCCACCAAGCGCTTCCGACGCTAGAGGT 311  
:  
98 SerGlyLeuTyraLacCysSerValLeuLysAsnSerHisPhePheIsva 114  
312 CAATACACACCCCTGCTGTGATGATGATATGACCCAGGAAAGAGAGCC 361  
||| |||  
114 IAsnValThr..... 117  
362 TGGGCGCCGACAGCTCTGTGGGGGTCAAGAGAC.....CCGCGCAGC 405  
||||| :  
118 ....GluAlaSerSerSerGlyAspAspArgLysAspAsnAspGlySer 132  
406 CAGCAGTGGGACAGA.....CCGCGCTTACACA 434  
:  
133 GluAspPheThrAsnAspAsnAsnIleArgAlaProTyrTrpThrAs 149  
435 GCGCTCAAGATGAGCGCGCGGTGATGCGACGCGCGGGGTACTGCTCG 484  
:  
149 ThrGluLysMetGluLysLysLeuHisAlaValSerAlaAlaAsnThrV 166  
485 TCGCGCTCAAGTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 534  
:  
166 alLysLeuArgCysProAlaArg...GluProHisProSerAsnGlyTrp 181  
535 ATGAAGACAGCAGCGCGCTTACG.....CGCCAGAGCGCGGCTGA 575  
:  
182 LeuLysAsnGlyLysGluPheLysGlnGluHisAspGlyLeuGlyTyr 198  
576 GCGCCGAAAGAAAGATGGAGCAGCTGAGCCTGAAGCAAGTGGCGGAGG 625  
:  
198 sValArgAsnGlnHisTrpSerLeuIleMetCysLeuValValProseA 215  
626 ACAGCGCAAAATACACTGCGCGCTGTGCAACCGCGCGGCGCGCATCAAC 675  
:  
215 sPlyGlyLLeuTyrThrCysIleValGluAsnGlnHisGlySerIleasn 231  
676 GCGACCTACAGGTGATGTATCCAGCGGACCGCTCCAGCGCGGTGCT 725  
||||| :  
232 HisTrpTyrHisLeuAspValIleGluLysSerSerHisArgProIle 248  
726 CACAGCGACGACCGCGGAAACAGCAGGTGAGCTTGGGGGAGACACAGT 775  
| :  
248 uGlnIleGlyLeuProAlaAsnThrThrAlaValAlaGlyLysAspAlaG 265  
776 CCTTCAGTCAAGGTGGCGCACGACGATGAACCGCGGTATCCAGTGGCTG 825  
:  
265 IurPheValCysLysValTyrSerAspAlaGlnProHisIleLeuGlyTrpAl 281  
826 AAGCGCTGTGAGTACGCGCGCGGCGGCGGCGGCGGCGGCGGCGGAGT 875  
:  
282 ArgTyrIleGlu.....LysAsnGlySerArgPheGlyIva 293  
876 GGGCGGC...CAGAGTTTGTGTGCTGCCCGCGGTGACGTGTGCTGCG 922  
| ||| ||| ||| :  
293 IAspGlyLeuProTyrPheLysValLeuLysAlaIleGlyVal.....A 308  
923 GCGCGGAGCGCTTACCTCAATTAAGCTCTCATCATCAGCGCGCGGCGAG 972  
:  
308 snValThrAspGluGluIleGluValLeuTyrValAlaArgAsnValSerPhe 324







```

244 GlnLeuAspValValGluArgSerProHisArgProIleLeuGlnAlaG1 260
735 GCAACCCGTCAGACAGAGCGTGAATTGGGGGGACCACTCTTCAGT 784
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
260 YleProAlaAsnLysThrValAlaLeuGlySerAsnValGluPheMetC 277
785 GCAAGGTGGCCGACGAGCTGATCCAGTGGCTGAAGCGCGT 834
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
277 YstYsValYtYSerAspProGlnProHisIleGlnTrpLeuYshIstIle 293
835 GAGTACGCGCCGAGGCGCCCAACTCCACCATCATGATGGGCGGCCA 884
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
294 GluVal.....AsnGlySerLysIleGlyProAspAsnLe 305
885 GAAGTTTGTG...GTGCTGCCACGAGGTGAGCTGTGCTGGCGCCGACG 931
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
305 UProTyValGlnIleLeuLysThrAlaGlyVal.....AsnThrThra 320
932 GCTCTTACCTCAATAGCTGCTCATCACCGCTGCCGACGAGCATGCG 981
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
320 spLysGluMetGluValLeuHisLeuArgAsnValSerPheGluAspAla 336
982 GGCATGTACATTCGCTTGGGCCAACAACATGAGGTACAGTTCGCGAG 1031
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
337 GlyGluTrpThrCysLeuAlaGlyAsnSerIleGlyLeuSerHisHisSe 353
1032 CGGCTTCTCTACCGCTGCTGCCA.....GACCCAAACCGCCAGGCGCAC 1075
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
353 rAlaTrpLeuThrValLeuGlnAlaLeuGluGluArgPro..... 366
1076 CTGTGGCTCTGCTGCTGCTGGCCACTAGCCTGGCGTGGCGGTGCATC 1125
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
367 .....AlaValMetThrSerProLeuTrpLeuGluIleIle 378
1126 GGCATCCACGCGCGCTGCTTCATC.....CTGGGCAACCCGCT 1166
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
379 IleTyCysThrIleAlaPheLeuIleSerCysMetValGlyValIle 395
1167 CCGTGTGGCTTGGCCAGCGCCAGAAAGACCG..... 1197
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
395 eValTyLysMetLysSerGlyThrLysLysSerAspPheHisSerGlnM 412
1198 ..TGCACCCCGCGCGCTGCCCTGCCCTGGGACCGCCGCGCGGG 1245
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
412 eValAlaValHisLysLeuAlaLysSerIleProLeuArgAlaValThr 428
1246 ACGCGCGCGACCGCGAGGAGACAAGAC.....CT 1277
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
429 ValSerAlaAspSerSerAlaSerMetAsnSerGlyValLeuLeuValAr 445
1278 TCCCTGCTTGGCGCGCTCAGCGCTGAGCCCTGTGGGCTGTGAGG 1327
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
445 gProSerArgLeuSerSerSerGlyThrProMetLeuAlaGlyValSerG 462
1328 AGCATGGGTCTCGGCGACGCCCGCCAG 1353
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
462 LutyrgLeuLeuProGluAspProArg 470
seq_name: SwissProt_39:FGRL_RAT
seq_documentation_block:
ID FGRL_RAT STANDARD; PRT; 822 AA.
AC 004589;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE BASIC FIBROBLAST GROWTH FACTOR RECEPTOR 1 PRECURSOR (BFGF-R)
DE (EC 2.7.1.112) (MFR).
GN FGFR1 OR FLG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OK NCBI_TaxId=10116;

```

```

RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=93176824; PubMed=8382532;
RA Yazaki N., Hiroko F., Mitsuhiro O., Toshisuke K., Nobuyuki I.;
RT "The structure and expression of the FGF receptor-1 mRNA isoforms in
   rat tissues."
RL Biochim. Biophys. Acta 1172:37-42(1993).
CC -1- FUNCTION: RECEPTOR FOR BASIC FIBROBLAST GROWTH FACTOR. A SHORTER
CC FORM OF THE RECEPTOR COULD BE A RECEPTOR FOR ACIDIC FGF (aFGF).
CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: D12498; BAA02059.1; -.
DR PIR: S29840; S29840.
DR HSSP: P11362; IRegI.
DR InterPro: IPR000719; -.
DR InterPro: IPR001245; -.
DR InterPro: IPR003006; -.
DR Pfam: PF00067; Ig_3.
DR Pfam: PF00069; Kinase; 1.
DR PRINTS: PS00109; TYRKINASE.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP_1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR_1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM_1.
DR Receptor: Glycoprotein; Tyrosine-protein kinase: ATP-binding;
KW Transferase; Phosphorylation; Transmembrane; Signal;
KM Immunoglobulin domain.
FT SIGNAL 1 21
FT CHAIN 22 822
FT DOMAIN 22 376
FT TRANSMEM 377 397
FT DOMAIN 398 822
FT DOMAIN 48 108
FT DOMAIN 171 237
FT DOMAIN 270 348
FT DOMAIN 478 767
FT NP_BIND 484 492
FT BINDING 514 514
FT ACT_SITE 623 623
FT MOD_RES 654 654
FT DISULFID 55 101
FT DISULFID 178 230
FT DISULFID 277 341
FT CARBOHYD 77 77
FT CARBOHYD 117 117
FT CARBOHYD 227 227
FT CARBOHYD 240 240
FT CARBOHYD 264 264
FT CARBOHYD 296 296
FT CARBOHYD 317 317
FT CARBOHYD 330 330
SQ SEQUENCE 822 AA; 91824 MW; E59D924D0A1DE5C5 CRC64;

alignment_scores:
Quality: 423.50 Length: 443
Ratio: 1.586 Gaps: 14
Percent Similarity: 60.271 Percent Identity: 27.540

alignment_block:
US-09-598-042a-2 x FGRL_RAT ..
Align seg 1/1 to: FGRL_RAT from: 1 to: 822

```

```

979 GGGGCGCATATACATCTGCCTTGAGCCCAACACCATTGGCTACAGCTTCG 1028
      |||||   |||   |||||:::|||||   |||   :
336 AAlaglyglutyrThrcysleuAlaglyAsnserIleglyLeuSerHisH1 352
      ::|||:::|||||
1029 CAGCGCCTTGTCCACCGGTGTGCCA.....GACCCAAACCGCCAGGCG 1072
      :|||||:::|||||
352 sSerAlatrPleuthrValleugluAlaleugluArgPro..... 366
      :|||||:::|||||
1073 CACCTGTGGCTCTCTGCTGCGCACACTAGCCGTGCGCGCGTGTGC 1122
      ::|||    |||||    ::|||
367 .....AlaValMetThrSerProleuTyrlleugluile 377
      :|||||
1123 ATGCGCATCCCAAGCGGCGCTGTCTTATC.....CTGGCACCT 1163
      |||   |||||   ::|||
378 IlelleTyrcysthrdylAlaphleuIleSerCysmetValcylSerVa 394
      :|||||:::|||||
1164 GCCTCGTGTGGCTTTGGCCAGCCCAAGAAGCG..... 1197
      :|||||:::|||||
394 llelleTyrlYsmetylsSerglYThrylsySsrAsprHeHisSerg 411
      :|||||
1198 ....TGACCCCCCGCGCTGCCCCTCCCTCGCTGGGACCGCCGCCG 1242
      :|||||   |||   ::|||
411 lImetalavalHIslysleuAlalysSerlleProleuNrgArgluVal 427
      :|||||
1243 GGAGCGCGCCGCGACCGCAGCGAGACAAGAC..... 1275
      :|||||   |||   |||:::|:::
428 ThrValSerAlaspSerSerAlasermetasenSerglyValleuLeuVa 444
      :|||||   ::|||
1276 .CTTCCCTGTGGCGCGCCCTCAGCCCTGCGCTGGTGTGGGGCTGTG 1324
      :|||||   ::|||
444 lArgProSerArgyleuSerSerSerglyThrProMetleuAlaglyValS 461
      :|||||
1325 AGGACGATGCTGTCCGCGACGCCCGCCAG 1353
      ::|||:::   |||   |||:::
461 erglutyrgluLeuProgluasproArg 470
      :|||||

seq_name: Swissprot_39:BFR2_HUMAN

seq_documentation_block:
ID BFR2_HUMAN STANDARD; PRT; 654 AA.
DC Q01742:
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE FIBROBLAST GROWTH FACTOR RECEPTOR BFR-2 PRECURSOR (EC 2.7.1.112).
GN BFR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP MEDLINE=91274356; PubMed=1647213;
RX Seno M., Sasada R., Watanabe T., Ishimaru K., Igarashi K.;
RA "Two cDNAs encoding novel human FGF receptor."
RL Biochim. Biophys. Acta 1085:244-246(1991).
RT -FUNCTION: RECEPTOR FOR ACIDIC AND BASIC FIBROBLAST GROWTH
CC FACTORS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: X51619.1; NOT_ANNOTATED_CDS.
CC HSSP: P11362.1;IGI

```

DR InterPro: IPR000719; ..  
 DR InterPro: IPR001245; ..  
 DR InterPro: IPR003006; ..  
 DR Pfam: PF00047; 1g; 2.  
 DR Pfam: PF00069; pkinase; 1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR Receptor: Glycoprotein; Tyrosine-protein kinase; ATP-binding;  
 KW Transferase; Phosphorylation; Transmembrane; Immunoglobulin domain;  
 KW Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 654 FIBROBLAST GROWTH FACTOR RECEPTOR BFR-2.  
 FT DOMAIN 22 263 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 264 284 POTENTIAL.  
 FT DOMAIN 285 654 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 57 123 IG-LIKE DOMAIN.  
 FT DOMAIN 156 234 IG-LIKE DOMAIN.  
 FT DOMAIN 367 647 PROTEIN KINASE.  
 FT NP\_BIND 373 381 ATP (BY SIMILARITY).  
 FT BINDING 403 403 ATP (BY SIMILARITY).  
 FT ACT\_SITE 512 512 BY SIMILARITY.  
 FT MOD\_RES 543 543 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 113 113 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 126 126 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 150 150 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 182 182 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 203 203 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 214 214 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 235 235 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 654 AA; 73594 MW; F4954E7DC70FD233 CRC64;

alignment\_scores:  
 Quality: 379.50 Length: 329  
 Ratio: 1.956 Gaps: 7  
 Percent Similarity: 58.967 Percent Identity: 29.483

alignment\_block:  
 US-09-598-042a-2 x BFR2\_HUMAN ..

Align seg 1/1 to: BFR2\_HUMAN from: 1 to: 654

265 GCGCTGACGTCGTGCAAGGCGCAGCGCTGGAGCGCTGAGCGTC.. 312  
 ||| .....  
 5 G1yArpHe11eCyLeuValVal1Thre1a1Thr1eUsSer1eUa1 21  
 313 .....ACTACACCCCTCGTGTGCTGGATGACATTAGCCGAGGAAG 355  
 21 aArgProSerPheSerLeuVal..... 28  
 356 AGAGCCTGGGCGGACAGCTCTCTGTGGGGTCAAGAGACCCGCCAGC 405  
 29 .....GluAspThr1ThrLeu 33  
 406 CAGCAGTGGGACGACGCGCTTCACACAGCCCTCAAGATGAGGCGCG 455  
 : : : : :  
 34 GluPProGluG1yAlaProTy1TriPhraSnThr1eUySmetG1y1sAr 50  
 456 GGTGATCGCAGCGCGCGTGGTGTAGCTCGTGGGCTCAAGTGCCTGCA 505  
 : : : : :  
 50 GluH1sAlaValPro1a1a1aSnThVal1yGpHeaRyGysPro1a1aG 67  
 67 1yG1yAsnProMeProThrMetArg1rPleUySAsnG1y1ySgluPhe 83  
 556 AGC.....CGCCAGAGCGCGCTGAGCCGAGAGAGAGAGAGTGGAC 596  
 ||| : : : : :  
 84 1ySg1nG1n1uH1sArG11eG1yG1yTr1ySVal1ArGAsnG1n1H1sTrpSe 100  
 597 ACTGAGCCTGAAGAACCTGGGCGGAGACAGCGGCAAAATACACTGCC 646

seq\_name: SwissProt\_39:NTC3\_MOUSE  
 seq\_documentation\_block:  
 ID NTC3\_MOUSE STANDARD; PRT; 2318 AA.  
 AC Q61982;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE NEUROGENIC LOCUS NOTCH 3 PROTEIN.  
 GN NOTCH3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ICR X SWISS WEBSTER;  
 RC MEDLINE=95001556; PubMed=7918097;  
 RA Lardelli M., Dalstrand J., Lendahl U.;  
 RT "The novel Notch homologue mouse Notch 3 lacks specific epidermal  
 growth factor-repeats and is expressed in proliferating  
 neuroepithelium.";

```

RL Mech. Dev.46:123-136(1994).
CC -1- FUNCTION: NOTCH 1, 2 AND 3 PLAY A COMBINATIONAL ROLE DURING
CC VARIOUS CELL FATE DECISIONS AND MORPHOLOGICAL MOVEMENTS IN THE
CC DEVELOPING CNS AND PROBABLY OTHER REGIONS OF THE EMBRYO.
CC -1- TISSUE SPECIFICITY: PROLIFERATING NEUROEPITHELIUM..
CC -1- DEVELOPMENTAL STAGE: CNS DEVELOPMENT..
CC -1- SIMILARITY: CONTAINS 34 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation --
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC -----
DR EMBL; X74760; CAA52776.1; -.
DR HSSP; P00740; IIXA.
DR MCD; MG1:99460; Notch3.
DR InterPro; IPR000152; -.
DR InterPro; IPR000561; -.
DR InterPro; IPR000800; -.
DR InterPro; IPR001438; -.
DR InterPro; IPR001881; -.
DR InterPro; IPR002110; -.
DR Pfam; PF00008; EGF_34.
DR Pfam; PF000023; ank; 6.
DR Pfam; PF00066; notch; 3.
DR PRINTS; PS00010; EGFBLOOD.
DR PROSITE; PS50086; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 18.
DR PROSITE; PS00022; EGF_1; 33.
DR PROSITE; PS01186; EGF_2; 27.
DR PROSITE; PS01187; EGF_CA; 17.
KW Differentialation; Neurogenesis; Repeat; EGF-like domain; Transmembrane;
KV ANK repeat; Glycoprotein.
FT DOMAIN 1 1643 EXTRACELLULAR.
FT TRANSMEM 1644 1664 POTENTIAL.
FT DOMAIN 1665 2318 CYTOPLASMIC.
FT DOMAIN 2242 2261 PEPT.
FT DOMAIN 39 78 EGF-Like 1.
FT DOMAIN 120 119 EGF-Like 2.
FT DOMAIN 159 196 EGF-Like 3.
FT DOMAIN 198 235 EGF-Like 4.
FT DOMAIN 237 273 EGF-Like 5.
FT DOMAIN 275 313 EGF-Like 6.
FT DOMAIN 315 351 EGF-Like 7.
FT DOMAIN 352 390 EGF-Like 8.
FT DOMAIN 392 430 EGF-Like 9.
FT DOMAIN 432 468 EGF-Like 10.
FT DOMAIN 470 506 EGF-Like 11.
FT DOMAIN 508 544 EGF-Like 12.
FT DOMAIN 546 581 EGF-Like 13.
FT DOMAIN 583 619 EGF-Like 14.
FT DOMAIN 621 656 EGF-Like 15.
FT DOMAIN 658 694 EGF-Like 16.
FT DOMAIN 696 731 EGF-Like 17.
FT DOMAIN 735 771 EGF-Like 18.
FT DOMAIN 772 809 EGF-Like 19.
FT DOMAIN 811 848 EGF-Like 20.
FT DOMAIN 850 886 EGF-Like 21.
FT DOMAIN 888 923 EGF-Like 22.
FT DOMAIN 925 961 EGF-Like 23.
FT DOMAIN 963 999 EGF-Like 24.
FT DOMAIN 1001 1035 EGF-Like 25.
FT DOMAIN 1037 1083 EGF-Like 26.
FT DOMAIN 1085 1121 EGF-Like 27.
FT DOMAIN 1123 1159 EGF-Like 28.
FT DOMAIN 1161 1204 EGF-Like 29.
FT CALCULON-BINDING (POTENTIAL).

```

FT	DOMAIN	1206	1245	EGE-LIKE 31.
FT	DOMAIN	1247	1288	EGE-LIKE 32.
FT	DOMAIN	1290	1326	EGE-LIKE 33.
FT	REPEAT	1336	1374	EGE-LIKE 34.
FT	REPEAT	1388	1428	LN/NOTCH 1.
FT	REPEAT	1429	1467	LN/NOTCH 2.
FT	REPEAT	1468	1503	LN/NOTCH 3.
FT	REPEAT	1839	1868	ANK 1.
FT	REPEAT	1872	1902	ANK 2.
FT	REPEAT	1906	1935	ANK 3.
FT	REPEAT	1939	1968	ANK 4.
FT	REPEAT	1972	2001	ANK 5.
FT	DISUFID	43	55	BY SIMILARITY
FT	DISUFID	49	66	BY SIMILARITY
FT	DISUFID	68	77	BY SIMILARITY
FT	DISUFID	83	94	BY SIMILARITY
FT	DISUFID	88	107	BY SIMILARITY
FT	DISUFID	109	118	BY SIMILARITY
FT	DISUFID	124	135	BY SIMILARITY
FT	DISUFID	129	145	BY SIMILARITY
FT	DISUFID	147	156	BY SIMILARITY
FT	DISUFID	163	175	BY SIMILARITY
FT	DISUFID	169	184	BY SIMILARITY
FT	DISUFID	186	195	BY SIMILARITY
FT	DISUFID	202	213	BY SIMILARITY
FT	DISUFID	207	223	BY SIMILARITY
FT	DISUFID	225	234	BY SIMILARITY
FT	DISUFID	241	252	BY SIMILARITY
FT	DISUFID	246	261	BY SIMILARITY
FT	DISUFID	263	272	BY SIMILARITY
FT	DISUFID	279	292	BY SIMILARITY
FT	DISUFID	286	301	BY SIMILARITY
FT	DISUFID	303	312	BY SIMILARITY
FT	DISUFID	319	330	BY SIMILARITY
FT	DISUFID	324	339	BY SIMILARITY
FT	DISUFID	341	350	BY SIMILARITY
FT	DISUFID	356	367	BY SIMILARITY
FT	DISUFID	361	378	BY SIMILARITY
FT	DISUFID	380	389	BY SIMILARITY
FT	DISUFID	396	408	BY SIMILARITY
FT	DISUFID	403	419	BY SIMILARITY
FT	DISUFID	420	429	BY SIMILARITY
FT	DISUFID	436	447	BY SIMILARITY
FT	DISUFID	441	456	BY SIMILARITY
FT	DISUFID	458	467	BY SIMILARITY
FT	DISUFID	474	485	BY SIMILARITY
FT	DISUFID	479	494	BY SIMILARITY
FT	DISUFID	496	505	BY SIMILARITY
FT	DISUFID	512	523	BY SIMILARITY
FT	DISUFID	517	532	BY SIMILARITY
FT	DISUFID	534	543	BY SIMILARITY
FT	DISUFID	550	560	BY SIMILARITY
FT	DISUFID	559	569	BY SIMILARITY
FT	DISUFID	571	580	BY SIMILARITY
FT	DISUFID	587	598	BY SIMILARITY
FT	DISUFID	592	607	BY SIMILARITY
FT	DISUFID	609	618	BY SIMILARITY
FT	DISUFID	625	635	BY SIMILARITY
FT	DISUFID	630	644	BY SIMILARITY
FT	DISUFID	646	655	BY SIMILARITY
FT	DISUFID	662	673	BY SIMILARITY
FT	DISUFID	667	682	BY SIMILARITY
FT	DISUFID	684	693	BY SIMILARITY
FT	DISUFID	700	710	BY SIMILARITY
FT	DISUFID	705	719	BY SIMILARITY
FT	DISUFID	721	730	BY SIMILARITY
FT	DISUFID	739	750	BY SIMILARITY
FT	DISUFID	744	759	BY SIMILARITY
FT	DISUFID	761	770	BY SIMILARITY
FT	DISUFID	776	787	BY SIMILARITY
FT	DISUFID	781	797	BY SIMILARITY
FT	DISUFID	799	808	BY SIMILARITY
FT	DISUFID	815	827	BY SIMILARITY

```

FT DISULFID 821 836 BY SIMILARITY.
FT DISULFID 838 847 BY SIMILARITY.
FT DISULFID 854 865 BY SIMILARITY.
FT DISULFID 859 874 BY SIMILARITY.
FT DISULFID 876 885 BY SIMILARITY.
FT DISULFID 892 902 BY SIMILARITY.
FT DISULFID 897 911 BY SIMILARITY.
FT DISULFID 913 922 BY SIMILARITY.
FT DISULFID 929 940 BY SIMILARITY.
FT DISULFID 934 949 BY SIMILARITY.
FT DISULFID 951 960 BY SIMILARITY.
FT DISULFID 967 978 BY SIMILARITY.
FT DISULFID 972 987 BY SIMILARITY.
FT DISULFID 989 998 BY SIMILARITY.
FT DISULFID 1005 1016 BY SIMILARITY.
FT DISULFID 1010 1023 BY SIMILARITY.
FT DISULFID 1025 1034 BY SIMILARITY.
FT DISULFID 1041 1062 BY SIMILARITY.
FT DISULFID 1056 1071 BY SIMILARITY.
FT DISULFID 1073 1082 BY SIMILARITY.
FT DISULFID 1089 1100 BY SIMILARITY.
FT DISULFID 1094 1109 BY SIMILARITY.
FT DISULFID 1111 1120 BY SIMILARITY.
FT DISULFID 1127 1138 BY SIMILARITY.
FT DISULFID 1132 1147 BY SIMILARITY.
FT DISULFID 1149 1158 BY SIMILARITY.
FT DISULFID 1165 1183 BY SIMILARITY.
FT DISULFID 1177 1192 BY SIMILARITY.
FT DISULFID 1194 1203 BY SIMILARITY.
FT DISULFID 1210 1223 BY SIMILARITY.
FT DISULFID 1215 1233 BY SIMILARITY.
FT DISULFID 1235 1244 BY SIMILARITY.
FT DISULFID 1251 1262 BY SIMILARITY.
FT DISULFID 1256 1276 BY SIMILARITY.
FT DISULFID 1278 1287 BY SIMILARITY.

```

## alignment\_scores:

```

Quality: 314.50 Length: 1478
Ratio: 0.589 Gaps: 93
Percent Similarity: 36.130 Percent Identity: 22.192

```

## alignment\_block:

US-09-598-042a-2/rev x NTC3\_MOUSE ..

Align seg 1/1 to: NTC3\_MOUSE from: 1 to: 2318

```

3056 TGTGTGCTGGGAGCGGCTGTGTCATCAGCATGCGTGAAGACGAGCG 3007
    |||  ::  |||  |||  ::  |||  |||  ::  |||  |||  ::
163 CysArgSerGlyThrThrCys....ArgHisGlyGlyThrCysLeuAs 177
3006 GGTCCCGGGCCCGGAGAGACGACAGTCC..... 2976
177 nThpProGlySerPheArgCysGlnCysProLeuGlyThrThrCysLeuL 194
2975 ..TGC...AACCCAGCTTC.....CTCCATCA..... 2952
194 euCysGluAsnProValAlaProCysAlaProSerProCysArgAsnGly 210
2951 ...TTATTAATATATCTCTCATTTCTTAATATAA..... 2919
211 GlyThrCysArgGlnSerAspValThrThrCysCysAlaCysLeuPr 227
2918 ....TACCAAGGCCCC...TCTCTGTGTGTCAGGGGAGAAATGAGTGGG 2876
227 OGlyPheGlnIleGlnAsnGlyValAlaAspAspCysProGlyH 244
2875 ATGAGCACAATGAGCATGAGCTCTCAGCCTTCAGGCTTGGGGCTGCTG... 2829
    ::  |||  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::
244 tsArgCysLeuAsnGlyGlyThrCysValAlaSerGlyValAlaSerThrTyra 260
2828 .....TGCCCC..... 2823
261 CysGlnCysProGlyIleThrPheGlnPheCysThrGluAspValAs 277

```

```

2822 .....CAACCCAGCCCAACAGCAGTAGGGGACTCTGGGACCCCAAG 2780
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
277 pGluCysGlnLeuGlnPro..... 283
2779 CAGGTGGCAAAATAGCCGCCAAGGCGAGGCGAGAGCGGGATGAG 2730
    |||  ::  |||  |||  ::  |||  |||  ::  |||  |||  ::
284 .....AsnAlaCysHisAsnGlyGlyThrCysPheAsnLeu 296
2729 GCGGGGACTGAGCGCGGAGACAGGCGGAGAGTGGGGGATGACGT 2680
    ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::
297 GlyGlyHisSerCysValCys.....ValAsnGlyThrPheGln 309
2679 GGAGCAGGGAAGTCCCTCATCACTATGAGCCTCAGGACACGCTACTG 2630
    |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::
309 yGluSerCysSerGlnAsnIleAspAspCysAlaThrAlaValCysPhe 326
2629 CAGGCTTCAGGCGACACCTCCCAAAAGCAGTCAGTGTGCTGTC 2580
    ||  |||  |||  |||  |||  |||  |||  |||  |||  |||
326 tsGlyValaThrCysHisAspArgValAlaSerPheThrCysAlaCysPro 342
2579 GCGAGCATATCTGCACCTGTGTG.....GCATGTGTGCGCG 2542
    ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
343 MetGlyLysThrGlyLeuLeuCysHisLeuAspAspAlaCysValSer 359
2541 AAGTGTGTCGCCAGGCGCATATCTGATGTGCGTGTGTATCG 2492
    :  ::  |||  |||  |||  |||  |||  |||  |||  |||  |||
359 nProCysHisGluAspAlaIle.....CysAspThrAsnProValSerG 374
2491 GACAGCATATCTGCACGCTGTGTCATGTCCAGACAGCATATCTGTC 2445
    |||  ::  |||  |||  |||  |||  |||  |||  |||  |||
374 lyArgAlaIleCysThrCysProGlyPheThrGlyAlaCysAsp 390
2444 .....ACATGTGTGTGCCAGCAATATCTG..... 2418
391 GlnAspValAspGluCysSerIleGlyAlaAsnProCysGlnHisLeuL 407
2417 .ACGTGTGTGAGTGTGAGGAGCAGCATATCTGTGTGTGCCAGGACG 2369
    |||  ::  |||  |||  |||  |||  |||  |||  |||  |||
407 yArgCysValAsnThrGlnGlySer.PheLeuCysGlnCysGlyArgGly 423
2368 ATATCTGCG.....TGC.....GTGTGTGTGCCGGA... 2342
    ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
424 TyrThrGlyProArgCysGlnThrAspValAlaSerGlnCysLeuSerIlePr 440
2341 ....CAGCATATCTGTGCATGCGTGTGTGTGCCAGCAGCATGTGTGCG 2296
    ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
440 ocCysArgAsnGlnAlaThrCysLeuAspArgIleGlyGlnPheThrCysI 457
2295 TGTGT..... 2276
457 IecysMetAlaGlyPheThrGlyThrTyrcysGlnValAlaSerIleAspGlu 473
2275 ATA.....TCCGCGTGTGTGCCAGGAGCATATCTGCGCTGTGACGCTG 2232
474 CysGlnSerSerProCysValAlaSerGlyValCysLys..AspArgY 489
2231 TGTGTGGA.....AGTGTGTGTGCCAGGAGCATATCTGTGTGTGCTG 2188
    |||  ::  |||  |||  |||  |||  |||  |||  |||  |||
489 alaSerGlyPheSerCysThrCysProSerGlyPheSerGlySerMetCys 505
2187 GTGTCCAGACAGCATATCTGTGCACACGTG..... 2158
506 GlnLeuAspValAspGluCysAlaSerThrProCysArgAsnGlyAlaY 522
2157 .TGTGTGTGTGTGCCAGCAATATCTGTCATGTGTGTGTCAGG... 2113
    |||  |||  |||  |||  |||  |||  |||  |||  |||
522 scCysValAspGlnProAspGlyThrGluCysArgCysAlaGluGlyPheG 539
2113 ..... 2113
539 IuGlyThrLeuCysGluArgAsnValAlaSpAspCysSerProAspProCys 555

```

```

2112 CAGCATTAATCTGTGTCAG.....GCACATATCTGTGACGTGCT 2069
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
556 HisHisglArgCysValAspGlyLeuAlaSerPheSerCysAlaCysAl 572
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
2068 GTCCGGATACCATATCTGCAGTGTGTGCTCCAGACA..... 2031
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
572 aProGlyThrGlyLeuArgCysGluSer..GlnValAspGluCysArg 588
2030 GCATATCCGTGT..... 2019
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
589 SerLeuProGlySerArgGlyGlyGlyLeuSerCysLeuValAspLysTy 605
2018 .GTGTGTGTGTCCAGCAGCAGCATGTGCGCATGGCTGCT..... 1980
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
605 rLeuCysArgCysProProGlyThrGlyValAlaSerCysGluValAla 622
1979 .....GCGTGT.....GTATGTTCCAGCGAC 1959
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
622 leaSerPheCysAlaSerAsnProCysThrPheGlyValCysArgAspGly 638
1958 ATGTCTTTGATGTGTGCGATGTCTGTGCTGCTGCTGCTGCTGCTGCTG 1909
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
639 lLeaSerArgTyAsp.....CysValCysGlnProGlyPheTh 651
1908 AGCTTATCTGTGTGCGCGCGCATATCTGTGATG.....TGC 1868
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
651 rGlyProLeuCysAsnValGlnLeaSerGluCysAlaSerSerProCysG 668
1867 TGTGTGCGTACGTGTGCTGCGAGGACAGTGTGCGCGCATGTGTGTGTC 1818
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
668 lGlyGluGlySerCysValAspGlyGlnGlyPheHisCysLeuGly 684
1817 ATACATGATCCAGGTAGTGTGTGTGCTG.....TGTGT 1783
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
685 ProProGlySerLeuProProLeuCysLeuProAlaSerHisProCysAl 701
1782 G.....TGTGTCCAGGGGTATGCTTACACACAGACTGCTGGGG 1742
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
701 aHisLysProCysSerHisGlyValCys.....HisAspAlaProGlyG 716
1741 TGTGTGCGCATCTCTGCGCGCATGCGGCTGCTGCTGCTGCTGCTGCTG 1692
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
716 lPheHisArgCysValCysGluProGlyTrpSerGlyProArgCysSerGln 732
1691 GTCTCTCATCTCCAGACTGCTGTGTGCGCGCGCGCGCGCGCGCGCTG 1642
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
733 SerLeuAlaProAspAla..CysGluSer..GlnPro.....C 744
1641 GCATATACGTCGCGCTGACACTGATAGTGA..... 1609
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
744 lysGlnAlaGlyGly..ThrCysThrSerAspGlyLeuGlyPheArgCysTh 760
1608 .TGTGTGTGTGTA.....CCTTCC 1590
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
760 rCysAlaLarProGlyPheGlnGlyHisGlnCysGluValLeuSerProCys 777
1589 CTCACAGTGTAGTGTGTGTGAG..... 1567
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
777 hrPro.....SerLeuCysGlnHisGlyGlyHisCysGluSerAspPro 791
1566 .....AGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1530
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
792 AspArgLeuThrValCysSerCysProProGlyTrpGln..... 804
1529 TTTGGGGTACAATTAAGGCGAGCAACTGGGCGCTGGCGCGCGCGCGCT 1480
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
805 .....GlyProArgCysGlnGlnAspValAspGluCysA 816
1479 GGGGGGCTGCCGAGACCCATGCTCTTCACAC.....AGCGCG 1442
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
816 lArgLysAlaSer.....ProCysGlyProHisGlyThrCysThrAsnLeu 830
1441 CCGCAATACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1392

```

---

```

      ||||| ||||| ||||| ||||| ||||| ||||| |||||
831 ProGlyAsnPheArgCysIleCys..... 838
1391 AACTTAGGCGCAGCAACTGGGCTGGGCCAGTAAGTGTGGGGGCTGTC 1342
838 ..... 838
1341 CGGAGACCCATGCTCTTCACACAGCCCGCACACAGGCGCAGCGCTGAGG 1292
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
839 .....HisArgGlyTyThrGlyProPheCysAspGlnAspIleAsp 852
1291 CGGCGAAGGAGGAAAGCTTGTCTGCGGTGGCGGTGCGGGCGCTGCC 1242
853 ..AspCysAspProAsnProCysLeuHis..... 861
1241 GCGGCGCGGTGCCAGCGAGGAGGAGGCGCGCGCGGTGTCAGCGCTT 1192
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
862 .GlyGlySerCysGlnAspGlyValGly...SerPheSerCysSerCys. 876
1191 CTTCGGGCTGTGGCAAGCCACAGGAGCAGG.....GTGCCAGGATGA 1148
877 LeuAspGlyPheAlaGlyPro..ArgCysAlaArgAspValAspGluCysL 893
1147 AGACAGCGCGCGGTGGATGCGCATGACACAGGCGCGCGCGCATGATG 1098
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
893 euSerSerProCysGlyProGlyThrCysThrAspHis.....Val 906
1097 GCCGAGCAGCAGAGGCGCACAGGTGCGCGCTTGTGGTGTGCGCAG 1048
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
907 lAlaSerPheThrCysAlaCys.....ProProGlyTyGly..GlyPhe 920
1047 CACGGTGAGAAAGCGCTGCGGAAAGCTGTAGCCCA...TGTGTGTGCGCG 1001
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
920 eHisCysGlnIleAspLeu..ProAspCysSerProSerSerCysPheAsn 936
1000 CAAGCGCAATGT.....ACATGCCCGCATGCTGTC 972
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
937 GlyGlyThrCysValAspGlyValSerSerPheSerCysLeuCysAlaPr 953
971 TGGCGGCGCAGCGGTGATGAGCAGCTTATTGAGTGAAGCGCGTGGCGG 922
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
953 oGlyTyThrGly..... 957
921 CGACCACAGTCACCCGTGGCGAGCAGCCACAATCTTGGCGCGCCA... 875
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
958 .....ThrHisCysGlnTyGlnAlaAspProCysPheSerArgProCys 972
874 .....CATGATGCTGTGAGTGTGGCGG 852
973 LeuHisGlyGlyIleCysAsnProThrHisProGlyPheGluCys..... 987
851 CCTCTGCGCGCGGTACTCCAGCGGCTTACAGCAGTGAATCAGCGGCTTAC 802
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
988 .....ThrCysArgGlnGlyPheThrCylSerGlnCysGlnA 1000
801 GTCGC.....TGC.....GCACCTTGC..... 785
1000 snProValAspTrpCysSerGlnAlaProCysGlnAsnGlyGlyArgCys 1016
784 .....ACTGGAAGGAC 774
1017 ValGlnThrGlyAlaTyCysIleCysProProGlyTrpSerGlyArgLe 1033
773 GTGTGCGCGCGGAGGTCCAGCGGTGCTTCAAGGAGGCTGCGCGCTGAG 724
      : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
1033 uCysAspIleGlnSerLeuProCysThrGlnAlaAlaGlnMetGly 1050
723 CACGG..... 719
1050 aArgLeuGluGlnLeuCysGlnGlyGlyCysIleAspLysGly 1066
718 .....GCTTGAAGGCGTCCGTGATGATCAT..... 692

```

```

1067 ArgSerHisIleValCysProGluGlyArgThrGlySerHisCysGlu 1083
691 .....CCACCTTGTAAGTGGCGTTGATGGCCCGCGCGGTCCA 652
1083 uHisGluValAspProCysThrAlaGlnProCysGlnHisGlyThrC 1100
651 CACGGCGC..... 644
1100 ysArgGlyThrMetGlyGlyThrValCysGluCysProAlaGlyTyrAla 1116
644 ..... 644
1117 GlyAspSerCysGluAspAsnIleAspGluCysAlaSerGlnProCysGlu 1133
643 .....AGTGATTTGGCGGTTC 625
1133 nasnglyGlySerCysIleAspLeuValAlaArgTyrLeuCysSerCysP 1150
624 CTCGC..... 620
1150 roProGlyThrLeuGlyValLeuCysGluIleAsnGluAspAspCysAsp 1166
620 ..... 620
1167 LeuGlyProSerLeuAspSerGlyValGlnCysLeuHisAsnGlyThrCys 1183
619 .....GCCGACGTTCTTCACGGCTGACGTGCACCTTCTTCTTCT 580
1183 sValAspLeuValGlyGly..PheArgCysAsnCysPro.....Pro 1196
579 GGGCTCAGCGGCGCTGGCGCGCTCAAGGCGTG..... 547
1197 GlyTyrThrGlyLeuHisCysGluAlaAspIleAsnGluCysArgProGlu 1213
547 ..... 547
1213 yAlaCysHisAlaAlaHisThrArgAspCysLeuGlnAspProGlyGly 1230
546 .....GTCCGCTTCATCCACGATGTCGGCGCGGAGGTGCCCGCT 505
1230 IsPheArgCysValCysHisProGlyPheThrGlyProArgCysGlnIle 1246
504 GGCACGACGACCTTGAGCGC.....CACGGAGCTAC 476
1247 .....AlaLeuSerProCysGlnSerGlnProCysGlnHisGly..... 1259
475 CCACGGCGCGCTGCATCACCAGCGCTCATCTTGAGGCGTG..... 433
1260 .....GlyGlnCysArgHisSerLeuGlyArgGlyGlyLeuThrPhe 1275
432 .....TGTGA.....GCG 424
1275 hCysHisCysValProProPheThrGlyLeuArgCysGluArgValAla 1291
423 CGGTGCTGCCCC.....AC 411
1292 ArgSerCysArgGluLeuGlnCysProValGlyTyrLeuProCysGlnGln 1308
410 TGCT.....GGCTGGCGGGGCTCTTGG 388
1308 rAlaArgGlyProArgCysAlaCysProProGlyLeuSerGlyPro..... 1323
387 ACCCCAGAGAGAGTGGCG.....GCCCGAGGCTCTCTTC..... 350
1324 .....SerCysArgValSerArgAlaSerProSerGlyAlaThr 1336
349 .....CTGGCTAATGTCATCC 333
1337 AsnAlaSerCysAlaSerAlaProCysLeuHisGlyGlySerCysLeuP 1353
332 AGCAGCAGAGAGGTTGATTGACCTCAGGCTGCCAGACCGTTGG...T 286
1353 oValGlnSerValProPhePheArgCysValCysAlaProGlyTyrPglG 1370

```

```

285 GGCCTTGACACGATACACGCGGACATCTCCCGCTCCACCTTCACCT 236
1370 lYProArgCysGlnThrProSerAlaAlaProGluValProGluGluPro 1386
235 TCAGCCCTTGGCGACGA..... 218
1387 ArgCysProArgAlaAlaCysGlnAlaTysArgGlyAspGlnAsnCysAs 1403
217 .CGCGAAGCGGCTCCAGCGGCTGGATGTCGGC..... 182
1403 pArgGlyCysAsnThrProGlyCysGlyTyrPaspGlyGlyAspCysSer 1420
181 .....CATCCTTG.....TCACATGG..... 164
1420 euAsnValAspAspProTrrPargGlnCysGluAlaLeuGlnCysTrrParg 1436
163 .....TCAGCGGCGCGGCTCCCGCTCCACT 138
1437 LeuPheAsnAsnSerArgCysAspProAlaCysSerSerProAlaCysLe 1453
137 G.....GGCACTGCAGCGCGCATGTCGGCGC 112
1453 uTyrAspAsnPheAspCysTyrSerGlyValArgAspArgTrrCysAsn 1470
111 CAGCCGCGGCGACCT.....GCCGACGAATGTCGACCGCGGACGACA 68
1470 roValTyrGlyTyrCysAlaAspHisPheAlaAsp..... 1482
67 TCACGTGATGAAGACAATCCCGCGGCGCATGGCGCGGAGACATGCCA 18
1483 .....GlyArgCysAs 1486
17 CCGCGGCGGCGCAATTGC 2
1486 pGlnGlyCysAsnThr 1491
seq_name: SwissProt_39:PGBM_HUMAN
seq_documentation_block:
ID PGBM_HUMAN STANDARD: PRT: 4393 AA.
AC P98160; Q16287;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE BASEMENT MEMBRANE-SPECIFIC HEPARAN SULFATE PROTEOGLYCAN CORE
DE PROTEIN PRECURSOR (HSPG) (PERLECAN) (PLC).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92112994; PubMed=1730768;
RA Kallunki P., Tryggvason K.;
RT "Human basement membrane heparan sulfate proteoglycan core protein: a
RT 467-kD protein containing multiple domains resembling elements of the
RT low density lipoprotein receptor, laminin, neural cell adhesion
RT molecules, and epidermal growth factor.";
RL J. Cell Biol. 116:559-571(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin, and Colon;
RX MEDLINE=92235084; PubMed=1569102;
RA Murdoch A.D., Dodge G.R., Cohen I., Tuan R.S., Iozzo R.V.;
RT "Primary structure of the human heparan sulfate proteoglycan from
RT basement membrane (HSPG2/perlecan). A chimeric molecule with multiple
RT domains homologous to the low density lipoprotein receptor, laminin,
RT neural cell adhesion molecules, and epidermal growth factor.";
RL J. Biol. Chem. 267:8544-8557(1992).
RN [3]
RP SEQUENCE OF 1018-1472 FROM N.A.

```



```

RC TISSUE-COLON:
RX MEDLINE=91365376; PubMed=1679749;
RA Dodge G.R., Kovalevsky I., Chu M.L., Hassell J.R., McBride O.W.,
RA Y.H.F., Iozzo R.V.;
RT Heparan sulfate proteoglycan of human colon: partial molecular
RT cloning, cellular expression, and mapping of the gene (HSPG2) to the
RT short arm of human chromosome 1.;
RL Genomics 10:673-680(1991).
RN [4]
RP SEQUENCE OF 892-1398 FROM N.A.
RC TISSUE=Fibrosarcoma;
RX MEDLINE=92120660; PubMed=1685141;
RA Kallunki P., Eddy R.L., Byers M.G., Kestila M., Shows T.B.,
RA Tytgysaen K.;
RT Cloning of human heparan sulfate proteoglycan core protein,
RT assignment of the gene (HSPG2) to 1p36.1--p35 and identification of
RT a BamHI restriction fragment length polymorphism.;
RL Genomics 11:389-396(1991).
RN [5]
RP SEQUENCE OF 1-21 FROM N.A.
RX MEDLINE=94052171; PubMed=8234307;
RA Cohen I.R., Gressel S., Murdoch A.D., Iozzo R.V.;
RT "Structural characterization of the complete human perlecan gene and
RT its promoter.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:10404-10408(1993).
CC -1- FUNCTION: THIS PROTEIN IS AN INTEGRAL COMPONENT OF BASEMENT
CC MEMBRANES. IT IS RESPONSIBLE FOR THE FIXED NEGATIVE ELECTROSTATIC
CC CHARGE AND IS INVOLVED IN THE CHARGE-SELECTIVE ULTRAFILTRATION
CC PROPERTIES. IT INTERACTS WITH OTHER BASEMENT MEMBRANE COMPONENTS
CC SUCH AS LAMININ AND COLLAGEN TYPE IV AND SERVES AS AN ATTACHMENT
CC SUBSTRATE FOR CELLS.
CC -1- SUBUNIT: PURIFIED PERLECAN HAS A STRONG TENDENCY TO AGGREGATE IN
CC DIMERS OR STELLATE STRUCTURES.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES.
CC -1- PTM: CONTAINS THREE HEPARAN SULFATE CHAINS AS WELL AS N-LINKED
CC AND O-LINKED OLIGOSACCHARIDES.
CC -1- SIMILARITY: CONTAINS 4 LDL-RECEPTOR CLASS A DOMAINS.
CC -1- SIMILARITY: CONTAINS 10.5 LAMININ EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 LAMININ DOMAINS IV.
CC -1- SIMILARITY: CONTAINS 22 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 LAMININ G-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 4 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X62515; CAA44373.1; -
DR EMBL: M85289; AAA52700.1; -
DR EMBL: M64283; AAA52699.1; -
DR EMBL: S76436; AAB21121.1; -
DR EMBL: L22078; -; NOT_ANNOTATED_CDS.
DR HSSP: P00740; 11XA.
DR MIM: 142461; -
DR InterPro: IPR000034; -
DR InterPro: IPR000082; -
DR InterPro: IPR000561; -
DR InterPro: IPR001436; -
DR InterPro: IPR001791; -
DR InterPro: IPR002049; -
DR InterPro: IPR002172; -
DR InterPro: IPR003006; -
DR Pfam: PF000068; EGF_4;
DR Pfam: PF01390; SEA_1;
DR Pfam: PF00047; Ig; 22;
DR Pfam: PF00052; laminin_B; 3;
DR Pfam: PF00053; laminin_EGF; 8;
DR Pfam: PF00054; laminin_G; 3.

DR Pfam: PF00057; 1dL_recept_a; 4.
DR PRINTS; PR00010; EGFBLDOD.
DR PROSITE; PS00022; EGF_1; 9.
DR PROSITE; PS01186; EGF_2; 5.
DR PROSITE; PS01209; LDLRA_1; 4.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 11.
DR PROSITE; PS50068; LDLRA_2; 4.
KW Signal; Basement membrane; Proteoglycan; Repeat; Glycoprotein;
KW Heparan sulfate; Laminin EGF-like domain; Immunoglobulin domain;
KW Extracellular matrix; EGF-like domain.
FT SIGNAL 1 21
FT CHAIN 22 4393
FT 22 193
FT 194 404
FT 405 506
FT 507 1678
FT 1679 3688
FT 3689 4393
FT 197 236
FT 283 321
FT 323 361
FT 366 405
FT 405 506
FT 523 532
FT 533 732
FT 733 765
FT 766 815
FT 816 873
FT 876 935
FT 881 925
FT 926 1127
FT 936 1127
FT 1128 1160
FT 1161 1210
FT 1211 1267
FT 1271 1326
FT 1327 1336
FT 1337 1531
FT 1532 1564
FT 1565 1614
FT 1615 1672
FT 1679 1773
FT 1774 1867
FT 1868 1957
FT 1958 2053
FT 2054 2153
FT 2154 2246
FT 2247 2342
FT 2343 2438
FT 2439 2535
FT 2536 2631
FT 2632 2728
FT 2729 2828
FT 2829 2926
FT 2927 3023
FT 3024 3114
FT 3115 3213
FT 3214 3300
FT 3301 3401
FT 3402 3490
FT 3491 3576
FT 3577 3671
FT 3671 3701
FT 3701 3847
FT 3848 3883
FT 3886 3924
FT 3966 4104
FT 4106 4143
FT 4145 4178
FT 4243 4391

```



```

3250 LeuArgSerProLeuProTPrGlnHisArgLeuGlnGlyAspThrLeuI 3266
1199 GCACCCCGCGGCC.....TGCC 1215
3266 eileProArgValAlaGlnGlnAspSerGlyGlnTyrIleCysAsnAlaT 3283
1216 CTTCCCTGGCTGGGCAC.....CG 1235
3283 hrSerProAla.GlyHisAlaGlnAlaThrIleIleLeuHisValGluSe 3299
1236 CCCGCCGGGGAGCGCCCGCAGCCAGCGGAGACAGACCTTCCTCGT 1285
3299 rProProTyrAlaThr.....ThrValProGln 3309
1286 TGGCCCGCCCTGAGCCTGGCCCTGTGGGGCTG...TGTGAGAGCAT 1332
3309 IsAlaSerValGlnAlaGlyGlnThrValGlnLeuGlnCysLeuAlaHis 3325
1333 GGGTCTCG.....GCAGCCCG 1349
3326 GlyThrProProLeuThrPheGlnTPrSerArgValGlySerLeuPr 3342
1350 CCACACTTACTGGGCCCGCCAGCTGTGCTGCGCTAGT..... 1390
3342 oglyArgAlaThr..AlaArgsnGlnLeuLeuHisPheGlnArgAlaAla 3358
1391 .....GTACCCCAACTCTACACAGACATC 1416
3359 ProGlnAspSerGlyArgTyrArgCysArgValThrAsnLysValGlySe 3375
1417 CACACACACACACT.....GTATGCGGCGCGCTGTGGAGAC 1457
3375 rAlaGlnAlaPheAlaGlnLeuLeuValGlnGlyProPro.....GlyS 3390
1458 CATGGCTCTCGCGCAGCCCGCCAGCACTTACTGGGCCCGCCAGCTGC 1507
3390 eileuProAlaThrSerIleProAlaGlySerThrProThrValGlnVal 3406
1508 TGGCCCTTAAGTTTACCCCAACTCTACACAGACATCCACACACACAC 1557
3407 ThrProGlnLeuGlnThrLysSerIleGlyAlaSerValGlnPheHisCys 3423
1558 ACACACACTCTCACACACATCACAGT.....GGAG 1589
3423 sAlaValProSerAspArgGlyThrGlnLeuArgTPrPheLysGlnGlyG 3440
1590 GGCAGAGTCCACACACATCTACTATCATGTAGACGGCAGCCGTATCT 1639
3440 LysGlnLeuProProGlyHis.....SerValGlnAspGlyValLeuArg 3454
1640 GCAGAGGCGCAGGGGGCGCGCCAGACAGCAGCACTGGAGATGGAGG 1689
3455 IleGlnAsnLeu.....As 3459
1690 ACGGAGCTGCAGAGAGGAGGAGCC.....CATGGCG 1724
3459 pGlnSerCys.....GlnGlyThrTyrIleCysGlnAlaHisGly. 3472
1725 AGGAGGATGGCCAGCAGCCCGCAGCA..... 1750
3473 .....ProTPrGlyLysAlaGlnAlaSerAlaGlnLeuValIleGlnAla 3487
1751 .....GTCTGTGTGAGGACATAGCCCG 1773
3488 LeuProSerValLeuIleAsnIleArgThrSerValGlnThrValValVa 3504
1774 TGGACACACA.....CACACACACACAC 1796
3504 IGIyHisAlaValaGlnPheGlnCysLeuAlaLeuGlyAspProLysPrpG 3521
1797 ACACCTACCTGATGATGTATATGCACACACATGCGCGACAGTGTGCTGCT 1846
3521 InValThrTPrSerLysValaGlyLysHisLeuAlaTPrpGlyIleValaGln 3537

```

```

1847 GAAGGCACAGCTACGACACACACATGCACAGATATGCCGCTGGGCA 1896
3538 SerGly.GlyValValArgIleAlaHisValGlnLeuAlaSpAlaGly. 3553
1897 CACAGATTAAGCTGCCAAATGCAGCAGCAGCAGACATGCGCAGAC 1946
3554 ..GlnTyrArgCys.....ThrAlaThrAsnAlaIleGlyThr 3565
1947 ATACAGAGCAT 1958
3566 ThrGlnSerHis 3569

seq_name: SwissProt_39:PTPD_HUMAN
seq_documentation_block:
ID PTPD_HUMAN STANDARD; PRT; 1912 AA.
AC P23468;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROTEIN-TYROSINE PHOSPHATASE DELTA PRECURSOR (EC 3.1.3.48) (R-PTP-
DE DELTA).
GN PTPRD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND MUTAGENESIS OF ARG-1178.
RX MEDLINE=95204468; PubMed=7896816;
RA Pulido R., Krueger N.X., Serra-Pages C., Saito H., Streuli M.;
RT "Molecular characterization of the human transmembrane protein-
RT tyrosine phosphatase delta. Evidence for tissue-specific expression of
RT alternative human transmembrane protein-tyrosine phosphatase delta
RT isoforms."
RL J. Biol. Chem. 270:6722-6728(1995).
RN [2]
RP SEQUENCE OF 390-1912 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91006018; PubMed=2170109;
RA Krueger N.X., Streuli M., Saito H.;
RT "Structural diversity and evolution of human receptor-like protein
RT tyrosine phosphatases."
RL EMBO J. 9:3241-3252(1990).
CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
CC PROTEIN TYROSINE + ORTHOPHOSPHATE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: DIFFERENT ISOFORMS ARE FOUND IN DIFFERENT
CC TISSUES DUE TO ALTERNATIVE SPLICING.
CC -1- PTM: A CLEAVAGE OCCURS THAT SEPARATES THE EXTRACELLULAR DOMAIN
CC FROM THE TRANSMEMBRANE SEGMENT.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 8 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L38929; AAC41749.1; -
DR EMBL: X54133; CAA38068.1; -
DR PIR: S12052; S12052.
DR HSSP: P18052; IYFO.
DR MIM: 601598; -.
DR InterPro: IPR000242; -.
DR InterPro: IPR000387; -.
DR InterPro: IPR001777; -.
DR InterPro: IPR003006; -.

```



```

353 TyrTyrIleIleIleGlnHisLysProLysAsnSerIleGluLeuTyrLysG1 369
1165 CTCCTGTGGCTTTGCCAGCCAGAGAGAGCCGTCA..... 1201
369 u1leAspIyValAlaThrThrArgTyrSerValAlaGlyLeuSerProT 386
1202 .....CCCCGGCGCTGCCCTCCCTCCCTGGGACCGCCCGCGGGGA 1246
386 YrSerAspTyrGluPheArgValAlaValAlaValAsnIleGlyArgG1c 402
1247 CGGCGCCGCGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1296
403 ProProSerGluProValLeuThrGlnThrSerGluGlnAlaProSerSe 419
1297 AGCGGTGGCCCTGTGTGGGCTGTGTAGAGAGAGAGAGAGAGAGAGAG 1346
419 rAla.....ProArgAspValGlnAlaArgMetL 429
1347 CCCCAGGACAT.....TACTGGGCCA..... 1368
429 euSerSerThrThrIleLeuValGlnTyrLysGluProGluGluProAsn 445
1369 GGGCCAGTGTGGCTGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 1418
446 GlyGlnIleGlnGlyTyrArgValTyrTyrThrMetAspProThrGlnH1 462
1419 CACACACACA.....CACTGTATTGGCGCGCCCTGTGTAGAGAGACA 1459
462 sValAsnAsnTrpMetLysHisAsnValAlaAspSerGlnIleThrTr1 479
1460 TGGGTCTCCGAGAGCCCGCCCACTTACTGTGGCCAGCCAGTGTGTG 1509
479 lGlyAsnLeuValProGlnLysThrTyr.....SerValLysValLeu 493
1510 GCCCTAAGTTATACC.....CCAAACTGTACAGACATACAGACACA 1550
494 AlapherThrSerIleGlyAspGlyProLeuSerSerAspIleGlnVal11 510
1551 CACACACACACACACTTCTACACACACTCAGACGTGAGGCGAGAGTCCA 1600
510 eThrGlnThrGlyValProGlnLysProLeuAsnPheLysIleGluProG 527
1601 CCAGCACATTCACATATCAGTGTAGAGCGCACCGATATCGAGAGGGCAC 1650
527 lueSerGluThrSerIle..LeuLeuSerTrpThr..... 537
1651 GGGGGGGCGCGCAGACAG.....GCAGACTGGGAGGATGAGAGGAC 1691
538 .....ProProArgSerAspThrIleAlaAsnTyrGlu..... 548
1692 GGAGCTGCAGAGAGAGGAGGAGCCATGCGAGAGAGGAA 1732
549 .....LeuValTyrLysAspIyLysGlnIleGluGln 560

```

seq.name: SwissProt\_39::LAR\_DROME

seq\_documentation\_block:

```

ID LAR_DROME STANDARD; PRT; 2029 AA.
AC p16621;
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROTEIN-TYROSINE PHOSPHATASE DLAR PRECURSOR (EC 3.1.3.48) (PROTEIN-
DE TYROSINE-PHOSPHATE PHOSPHOHYDROLASE).
GN LAR.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyroidae; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=90046860; PubMed=2554325;
RA Streuli M., Krueger N.X., Tsai A.Y.M., Saito H.;
RT "A family of receptor-linked protein tyrosine phosphatases in humans
RT and Drosophila."
RL Proc. Natl. Acad. Sci. U.S.A. 86:8698-8702(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON-S.
RX MEDLINE=96178473; PubMed=8596047;
RA Krueger N.X., van Vactor D., Wan H.I., Gelbart W.M., Goodman C.S.,
RA Saito H.;
RT "The transmembrane tyrosine phosphatase DLAR controls motor axon
RT guidance in Drosophila."
RL Cell 84:611-622(1996).
CC -1- FUNCTION: IT IS POSSIBLE THAT DLAR IS A CELL ADHESION RECEPTOR.
CC IT POSSESSES AN INTRINSIC PROTEIN TYROSINE PHOSPHATASE ACTIVITY
CC (PTPASE). IT CONTROLS MOTOR AXON GUIDANCE.
CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
CC PROTEIN TYROSINE + ORTHOPHOSPHATE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: SELECTIVELY EXPRESSED IN A SUBSET OF AXONS AND
CC PIONEER NEURONS IN THE EMBRYO.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 16 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M27700; AAA28668.1; -
DR EMBL; U36857; AAC47002.1; -
DR EMBL; U36849; AAC47002.1; JOINED.
DR EMBL; U36850; AAC47002.1; JOINED.
DR EMBL; U36851; AAC47002.1; JOINED.
DR EMBL; U36852; AAC47002.1; JOINED.
DR EMBL; U36853; AAC47002.1; JOINED.
DR EMBL; U36854; AAC47002.1; JOINED.
DR EMBL; U36855; AAC47002.1; JOINED.
DR EMBL; U36856; AAC47002.1; JOINED.
DR PIR; A36182; TDFELK.
DR HSSP; P28827; IRPM.
DR FlyBase; FBgn0000464; Lar.
DR InterPro; IPR000242; -
DR InterPro; IPR000387; -
DR InterPro; IPR001777; -
DR InterPro; IPR003006; -
DR Pfam; PF00102; Y-phosphatase; 2.
DR Pfam; PF00041; fn3; 9.
DR Pfam; PF00047; Ig; 3.
DR PRINTS; PR00014; ENTPEIIT.
DR PRINTS; PR00700; PRTYPHPTASE.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 2.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 2.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PP; 2.
KW Hydroxase; Receptor; Glycoprotein; Signal; Transmembrane;
KW Cell adhesion; Immunoglobulin domain; Duplication.
FT SIGNAL 1..32
FT CHAIN 33..2029 PROTEIN-TYROSINE PHOSPHATASE DLAR.
FT DOMAIN 33..1377 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1378..1402 POTENTIAL.
FT DOMAIN 1403..2029 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 50..118 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 154..216 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 249..308 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 320..417 FIBRONECTIN TYPE-III 1.
FT DOMAIN 418..512 FIBRONECTIN TYPE-III 2.
FT DOMAIN 513..607 FIBRONECTIN TYPE-III 3.
FT DOMAIN 608..706 FIBRONECTIN TYPE-III 4.

```

```

FT DOMAIN 707 809 FIBRONECTIN TYPE-III 5.
FT DOMAIN 810 906 FIBRONECTIN TYPE-III 6.
FT DOMAIN 907 1007 FIBRONECTIN TYPE-III 7.
FT DOMAIN 1008 1102 FIBRONECTIN TYPE-III 8.
FT DOMAIN 1103 1207 FIBRONECTIN TYPE-III 9.
FT DOMAIN 1492 1738 PROTEIN-TYROSINE PHOSPHATASE.
FT DOMAIN 1781 2029 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 1670 1670 BY SIMILARITY.
FT ACT_SITE 1961 1961 POTENTIAL.
FT DISULFID 57 111 POTENTIAL.
FT DISULFID 161 209 POTENTIAL.
FT DISULFID 256 301 POTENTIAL.
FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 553 553 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 666 666 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 721 721 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 774 774 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 915 915 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 962 962 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1183 1183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1304 1304 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 2029 AA; 229027 MW; 536A0C794D3DC800 CRC64;

```

```

alignment_scores:
  Quality: 299.00 Length: 500
  Ratio: 1.256 Gaps: 24
  Percent Similarity: 47.600 Percent Identity: 25.400

```

alignment\_block:  
US-09-598-042a-2 x LAR\_DROME ..

Align seq 1/1 to: LAR\_DROME from: 1 to: 2029

```

82 CGAGGATTGTGGCGAGGTGGCCGGCGGCGACATGCGCGTCA 131
   |||::: |||::: |||::: |||::: |||::: |||:::
40 ArglysProGlnasnGlnGlyValArgValGlyValAlaSerPheTy 56
132 GTGCCGATGAGGAGGAGCCCGCGCGTGCATGTGACCAAGATG 181
   |||::: |||::: |||::: |||::: |||::: |||:::
56 rcysAlaAlaArglyAspPropProserIleValITrpArglyasnG 73
182 GCCGACCATCCACAGCGCGTGGCGCTTCCGCGTCTG.....CCG 225
   |||::: |||::: |||::: |||::: |||::: |||:::
73 IyLysLysValSerGlyThrGlnSerArgTrpThrValLeuGlnGlnPro 89
226 CAGGGGCGAAGTGAAGCAGGTGAG.....CGGAGAGA 260
   |||::: |||::: |||::: |||::: |||::: |||:::
90 GlylyIleSerIleLeuArgIleGlnProValArgAlaGlyArgAspAs 106
261 TGCCGGCGTGTACGTGTGCAAGGCCACCAAGCGCTTGGC...AGCCTGA 307
   |||::: |||::: |||::: |||::: |||::: |||:::
106 Pala...ProTrpGlnCysValAlaGlnasnGlyValGlyAspAlaValS 122
308 GCGTCACACTACACCTGCTGCTGCTGATGATCAATTAGCCCAAGGAAGAG 357
   |||::: |||::: |||::: |||::: |||::: |||:::
122 eraLaAspAlaThrLeuThrIleTyGlu..... 131
358 AGCCTGGGGCGCCGACAGTCTCTGGGGGTCAAGAGACCCCGCCAGCA 407
   |||::: |||::: |||::: |||::: |||::: |||:::
132 .....GlyAspLysThrProAla..... 137
408 GCAGTGGGACAGCCGCGCTTACACAG...CCCTCAAGATGAGGGCGC 454
   ::|||::: |||::: |||::: |||::: |||::: |||:::
138 .....GlyPheProValIleThrGlnGlyProGly.....Thra 149
455 GGGTGTATGCACGGCCCGTGGGTAGCTCCGTGGCGCTCAAGTGGTGGCC 504
   |||::: |||::: |||::: |||::: |||::: |||:::
149 rgValIle.....GluValGlyIleThrValLeuMetThrCysLysAla 163

```

```

505 AGCGGCAACCCCTCGGCCCGACATCATGTGATGAAGAGCAGCAGCCTT 554
   |||::: |||::: |||::: |||::: |||::: |||:::
164 IleGlyAsnProThrProAsnIleTyTrpIleLysasnGln..... 177
555 GACGGCGCCAGAGCGCGTGGAGCCAGG.....AAGAAGAATGGACAC 598
   |||::: |||::: |||::: |||::: |||::: |||:::
178 .ThrLysValAspMetSerAsnProArgTySerLeuLysAspGlyPheL 194
599 TGAGCCTGAAGAACTCGGGCGCGGAGAGCAGGCGCAATATACCTCGCCG 648
   |||::: |||::: |||::: |||::: |||::: |||:::
194 euGlnIleGlnAsnSerArgGlnGlnAspGlnGlyTyGlnGlyVal 210
649 GTGTGCAACCGCGCGCGC.....GCCATCAAGCCCACTTA 683
   |||::: |||::: |||::: |||::: |||::: |||:::
211 AlaGlnAsnSerMetCylThrGlnIleSerLysAlaThrAsnLeuTyVa 227
684 CAAGTGTATGTATGATCAGCGGACCGCTTCCAAAGCCGCTCACAAGCA 733
   |||::: |||::: |||::: |||::: |||::: |||:::
227 LysValArgArgValProProThrPheSerArgProProGlnIle.... 242
734 CGCACCCTGTGAACAGCAGGTGACTTCGGGGGAGCACACGCTTCCAG 783
   |||::: |||::: |||::: |||::: |||::: |||:::
243 .....IleSerGluValMetLeuGlySerAsnLeuAsnLeuSer 255
784 TGCAAGTGTGCGAGCGAGCGTGAAGCCGATGATCATGAGTGCATGAAGCGCT 833
   |||::: |||::: |||::: |||::: |||::: |||:::
256 CysIleAlaValGlySerProMetProHisValLysTrpMetLys..... 270
834 GGAGTAGCGCGCCGAGGCGCCGACACATCCACATCATGTGGGCGCGC 883
   |||::: |||::: |||::: |||::: |||::: |||:::
271 .....GlySerGluAsp..... 274
884 AGAAGTTGTGTGTCGCCCGACGGGTGACGTGTGTCGCGCGCCGACGCG 933
   |||::: |||::: |||::: |||::: |||::: |||:::
275 .....LeuThrProGlnAsnGlnMet.....ProIleGly 284
934 TCCTACCTCAATAGCTGCTCATACCGCTGCCCGCCAGCAGCATGGCGG 983
   |||::: |||::: |||::: |||::: |||::: |||:::
285 ArgAsnValLeuGlnLeuIle.....AsnIleGlnIleSerAl 297
984 CATGTATCATCTGCTTGGCGCCACACCATGGGCTACCTCCGAGGG 1033
   |||::: |||::: |||::: |||::: |||::: |||:::
297 aaenTyTrpCysIleAlaAlaSerThrLeuGlnIleAspSerValS 314
1034 CTTTCTCCACCGCTGTCGACGACCAACCCGACGCGCACCTGTGGCC 1083
   |||::: |||::: |||::: |||::: |||::: |||:::
314 eraValValLysValGlnSerLeuProThrAlaProThrAspValGlnIle 330
1084 TCCTGCTCTCGGCGCACTAGCTGCGCGCGCGCTGATCGGCTATGCC 1133
   |||::: |||::: |||::: |||::: |||::: |||:::
331 SerGlnValThrAlaThrSerValArgLeuGlnIlePheSerTyTrpLysGlyP 347
1134 AGCGGCGCTGTCT.....TCATCCGGGCGACCTGCTCTGTGGC 1174
   |||::: |||::: |||::: |||::: |||::: |||:::
347 roGlnAspLeuGlnTyTrpValIleGlnIleTyTrpLysProLysAsnAlaAsn 363
1175 TTGCGCAGCCGACAGAAAGC..... 1195
   |||::: |||::: |||::: |||::: |||::: |||:::
364 GlnAlaPheSerIleIleSerGlyIleIleThrMetTyTrpValValAr 380
1195 ..... 1195
380 gAlaLeuSerProTyThrIleTyTrpGlnPheTyValIleAlaValAsnA 397
1196 .....CGTGCACCCCGCGCGCTGCCCTCCCGCTGGGCGACCGCGC 1237
   |||::: |||::: |||::: |||::: |||::: |||:::
397 snIleGlyArgGlyProProSerAlaProAlaThrLys..... 409
1238 CGCGGGGAGCGCCGCGACCGCAGCGAGAGACAAG.....ACCTTCCC 1281
   |||::: |||::: |||::: |||::: |||::: |||:::
410 .....ThrThrGlyGlnThrLysMetCylSerLysAlaP 420
1282 TCGT.....TGCGCGCCCTCAGCGCTGGCCCTGTGT 1313

```



```

460 .. ATGCGACGCCCC.....GTGGTAGCTCCGTGGCG 489
      ::::|
242 hELeuGlnrPrGoserAsnValIleAlaIleGluGlyAspAlaVal 258
      ::::|
490 CTCAGTGTGGTGGCCAGCGGCGCCGCGGACATCATCTGATGANA 539
      ::::|
259 LeuGluCysCysValSerIleYrPrPProSerPheThrTriPLeuAr 275
      ::::|
540 GGACGACGAGCCCTTGACGCGCCGCTGACGCCAGAGAGAGAGA 589
      ::::|
275 gEgLyLu.....GluValIleGlnLeuArgSerLysL 286
      ::::|
590 AGTGAGCA.....CTGAGCCTGAAGAACCGGGCGCG 621
      ::::|
286 ysrIserLeuLeuGlyGlySerAsnLeuLeuSerAsnValThrAsp 302
      ::::|
622 GAGGACAGCGGCAATATACACTGCGCGTGTGAMCCGCGCGGCCAT 671
      ::::|
303 AspAspSerGlyThrYrThrCysValValThrYrLysAsnGluAsnI 319
      ::::|
672 CAAGCCACCTACAGGTGATGTATCCAGCGGACCCGTTCCAGCCG 721
      ::::|
319 eSerAlaSerAlaGluLeuThrValLeu.....ValPrP 331
      ::::|
722 TGCTACAGGACGACGCGCGTGAACAGCGGTGAGCTTGGGGGAGAC 771
      ::::|
331 rGTrPheLeuAsnHisProSerAsnLeuYrAlaYrGluSerMetAsp 347
      ::::|
772 ACCTCTTCAGTGAAGTGGCGAGCGAGCGAGTGAACCGGTATCCAGT 821
      ::::|
348 IleGluPheGluCysAlaValSerGlyLysProValProThrValAsn 364
      ::::|
822 GCTGAAGCGGTGAGTACGGCGCCGAGGCGCCGACCACTCCACATCG 871
      ::::|
364 pMetLysAsnGlyAsp.....369
      ::::|
872 ATGAGGCGGCGCAGAGTTGTGTGTGCCACAGGAGTGCATGTGTG 921
      ::::|
370 .....ValValIleProSer..AspYrPheGln 378
      ::::|
922 CGGCGCGAGCGCTCTACTCAATAGCTGCTCATCCGCGCGCCCA 971
      ::::|
379 IleValIleGlySerAsnLeuArgIleLeuGlyValValLysSer 393
      ::::|
972 GGACGATGGCGCATGTACATGTGCTTGGCGGCAACCATGGCTACA 1021
      ::::|
394 ....AspGluGlyPheYrGlnCysValAlaGluAsnGluAlaGlyAsn 409
      ::::|
1022 GCTTCCGCGAGCGCTCTACCGGTGTGCGACAGCCAAACCGCCAGG 1071
      ::::|
409 IagInSerSerAlaGlnLeuIleVal.....ProLys 419
      ::::|
1072 CCACCTGTGGCTCTCTGCTGCGGCATAGCCCTGCGGCCGTGTGT 1121
      ::::|
420 ProAlaIleProSerSerIleLeuProSerAlaProArgAsp...Va 435
      ::::|
1122 CATGGGATCCACAGCGCGCTGTCTTCATCCTGGGACACCTGTCTGT 1171
      ::::|
435 IleuProValIleuValSerSerArgPheVal.....ArgLeuSer 449
      ::::|
1172 GGCCTTGCAGAGCCAGAGAACGCTGACACCCCGCGCTGCCCCCTCC 1221
      ::::|
449 rP.....449
      ::::|
1222 CTGCGTGGGACCGCGCGCG.....1242
      ::::|
450 .....ArgProAlaGluAlaValGlyAsnIleGlnThrPh 462
      ::::|
1243 .....GGAGCGCGCGCGAGCGAGCGAGAGACA 1270
      ::::|
462 eThrValPhePheSerArgGluGlyAspAsnArgGluArgAlaLeuAsn 479
      ::::|

```

```

1271 AGGACTTCCCTCGTGGCGCCCTCAGCGCT.....1302
      ::::|
479 hrThrGlnProGlySerLeuGlnLeuThrValIleGlyAsnLeuLysProGlu 495
      ::::|
1303 .....GGCCCTGGGT 1313
      ::::|
496 AlaMetYrThrPheArgValAlaAlaTyrAsnGluThrPglyProGly.. 511
      ::::|
1314 GGGGCTGTGTAGAGCATGGGTCTCCG.....GCAGCCCCCAGC 1354
      ::::|
512 .....GluSerSerGlnProIleLysValAlaThrGlnProG 524
      ::::|
1355 ACTTACTGGGCCAGGCGCCAGTGTGCGCTTAAGTTACCCAAACTC 1404
      ::::|
524 IuLeuGlnValProGlyProVal.....531
      ::::|
1405 TACACAGACATCCACACACACACTGTATGTGGCGCGCTGTGTAG 1454
      ::::|
532 .....GluAsnLeuHisAlaValSerThrSerProThrSe 543
      ::::|
1455 GAGCATGGGTCTCCGCGAGCCCGCAGCACTTACTGGCGCCAGCCAGT 1504
      ::::|
543 rIleLeuIleThrTrpGluProProAlaTyrAlaAsnGlyProValGln 560
      ::::|
1505 TGCTGGCCCTAAGTTGTATACCCCAACTCTACA 1536
      ::::|
560 LYTrArgLeuPheCysThrGluValSerThr 570
      ::::|
seq_name: SwissProt_39:BAR3_CHITE
seq_documentation_block:
ID BAR3_CHITE STANDARD: PRT; 1700 AA.
AC Q03376;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE BALBIANT RING PROTEIN 3 PRECURSOR.
DE
GN BR3.
OS Chironomus tentans (Midge).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomidae; Chironomidae; Chironominae; Chironomus.
OX NCBI_TaxID=7153;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=90172404; PubMed=1689777;
RA Paulsson G., Lendahl U., Galli J., Ericsson C., Wieslander L.;
RT "The Balbiant ring 3 gene in Chironomus tentans has a diverged
RT repetitive structure split by many introns."
CC J. Mol. Biol. 211:331-349(1990).
CC -! FUNCTION: USED BY THE LARVAE TO CONSTRUCT A SUPRAMOLECULAR
CC STRUCTURE. THE LARVAL TUBE. BALBIANT RING PROTEIN 3 COULD PLAY A
CC ROLE AS A TRANSPORT PROTEIN THAT BINDS TO OTHER PROTEINS
CC INTRACELLULARLY AND IN THE GLAND LUMEN IN ORDER TO PREVENT THESE
CC FROM FORMING WATER-INSOLUBLE FIBERS TOO EARLY.
CC -! SUBCELLULAR LOCATION: SECRETED.
CC -! TISSUE SPECIFICITY: SALIVARY GLAND.
CC -! DOMAIN: HAS 82 APPROXIMATE REPEATS OF CYS-X-CYS-X-CYS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL, X52263; CA36506.1; -.
CC PIR, S08167; S08167.
CC HSSP, P18055; 2MRB.
CC InterPro: IPR000853; -.
CC PRINTS, PR00876; MTNEMA00DE.

```



**alignment\_scores:**

Quality:	297.50	Length:	1007
Ratio:	0.785	Gaps:	58
Percent Similarity:	37.637	Percent Identity:	18.769

alignment\_block:

US-09-598-042A-2/rev x BAR3\_CHITE

Align seg 1/1 to: BAR3\_CHITE from: 1 to: 1700

```

2565 ACCTGCTGTGCATGTGTGTGCCGAAGTGTGTGCCACGACGACATATCTG 2516
      ||||| |||:::||||
752 ThrcyaspCysValCys..... ProGlnIynslnthrcy 763
2515 CATG.....TGTGCGTGGCTGTATTC 2493
      ||::: ||:::|||||:::|||||:::
763 sllleAlaProLysValTrpAspAlaLysThrcySerCylIleCysValA 780
2492 GGCACACATATCTGCACGCTGTGCATGTCCAGACACGATCTGTGCCAC 2443
      :: ||| ||::: ||:::
780 snProPro.....LysCysAsnSerProGlnValLeuLysAspPthr 793
2442 ATCTGTGTGTCAGGACGACATATCTGCACGCTGTGAGGTGAGGACGACAT 2393
      ||| ::: ||| ::: |||::: ||:::
794 CysCysCysGlyCysGlnAsnValLysSerCysLysAlaProGlnLysP 810
2392 TATC.....TGTGTGTGTGCCAGCAGCATATCTGCGTG 2358
      ||| ::: |||::: |||::: |||:::
810 eIIleGlnAsnIleCysAspCysAlaCysProAsnLysGlnCysLysA 827
2357 CG.....TGTGTGTCCGGACAGC 2338
      || || ||||| ||||| |||||
827 IaProLeuValTrpSerAspGluPheCysAspCysValCysProAsnSer 843
2337 ATATCTGTGCATGCGTGTGTGTCTCCGACAGCATGTCGTCGTGTGTG 2288
      ||::: ||::: ||::: ||| ::: |||
844 AlaSerMetLysThrcyLysLeuSerProLysGlnTrpAsnLysValThrcy 860
2287 TGACTTGACAGCATATCTCCGCTGTGTCCAGCAGCATATCTGCGCGCTGTG 2238
      | |||| |||
860 s.....ThrcyAspCysA 865
2237 CACGTGTGTGAGAGTGTGTGCCAGCAGCATATCTGCATGTG..... 2193
      ::: ||: ||||| ||::: ||:::
865 snProProLysProAspCysCysProGlnThrcyGlnLysTrpMetAspAs 881
2192 .....TGTGCGTGTCCAGACAC 2175
      ||::: ||||| ||::: ||:::
881 pLysCysLysCysGlnCysProAsnAlaGlnThrcyAspCysAlaGlyLyl 898
2174 .....ATATCTGCACACGCTGTGTGTGTGTGCA..... 2142
      ::::: ||||| || |||||
898 yslYspheAsnAspPheThrcySerCysCysLysProSerGlnLysLeu 914
2142 ..... 2142
915 AspCysThrGlyAsnThrLysTrpSerAlaGlnThrcyThrcyGlyLys 931
2141 .....GGCAATATCGT.....G 2129
      ||||| ||::: ||:::
931 sglYAspValAsnArgAsnCysGlyAsnLeuLysAsnPheAsnAspAsnL 948
2128 CATGTGTGTCTCAAGCAGCATATCTGTGTGCCAGCAGCATATCTG... 2082
      ||| ::: ||| ::: ||||| ||:::
948 euCysGlnCysGln.....CysLysAsnLysGlnGlnMetAla 960
2081 .....TGTCCACGCTGTGTG 2068

```

```

1171 ngluglulysCysglInCysglucysProLysaspLysProLysLysglInc 1188
1302 AGCGCTGAGGGCGGCCCAAGAGAGCTCTCTCCGCTGCGCTGCC 1253
1188 yspProgllyglInaspTrpAsnAsnHisglIncys...glncysgly... 1202
1252 GGGCGCGGCCCGCGGGCGGCTGCCAGGAGGAGGCGGCGGCGG... 1205
1203 CysPProthProAlaProthThrcysSerAsnAsnInLysTrSerAsnVa 1219
1204 .GGGTGACAGCGCTTCTTCTGGGCGCTGCAAAAGCCAGAGAGAGGTC 1156
1219 lserCysSerCysglYcysAsnProglLysProLysasn...gllycysb 1235
1155 CAGAGTGAAGA...CAGCGCGCGCTGGATGCCCATACACAGG 1115
1235 roglYasnInLlterpCysAspAsnThrcysArgCysValCysProLys 1251
1114 GCCACGCGAGCTAGTGGCGAGAGAGAGGAGCCACAGGTGGCCCTGGC 1065
1252 .....AsnMetglulysPro..... 1256
1064 GGTTTTGGGTCTGGCAGCAGGTGAGAGAGCGCTGCGAGCTGAGCC 1015
1257 .....AlaaspAsnCysLysr 1262
1014 CATGCTTGGCGCCAGCAGATGATGCCCGCATGCTGCTGGC... 968
1262 hrlYstrPTrpAsnaspLumetCysglInCysValCysLysProglYcys 1278
967 GGGCAGCGATGATAGCAGCTTATTGAGTAGAGAGCGCTGCGGCGCGAC 918
1279 ProglInglYcysglYalMetLysTrpAsnAlaAsnThrcys 1295
917 CACAGCTACCCG...TGGCAGCAGCACAACCTTGGCGCGCCAC 874
1295 rCysglucysProAlaaspLysAlaLysProAlaSerCysglYaspLysL 1312
873 ATCATGTG.....TGAGT 860
1312 ysserTrpAsnaspSerCysSerCysglInCysLysSerLysMetPro 1328
859 TGTGGCGGC...CCTGCGCGCGCT.....ACTCC 834
1329 CysglYcysProProAsnInglInTrpAsnInLysaspCysglucY 1345
833 ACAGCGCTTACACCTGATCA...CCGGCTTACAGCGCTGCCACCTT 787
1345 slyCysSerAlaThrlglYasnCysProAla..... 1355
786 GCACCTGGAAGAGCTGTGCCCGCGAGT...CCACCGTGTGT...TCA 743
1356 .....glYglInThrlTrpAsnSerInThrcysglInCysSerCysProAla 1370
742 CGGGGTCGCTCTGTGACACAGGAGCTTGAACGGTCCCTGATCA 693
1371 ThrglYcysCysThrlglYalaglInValTrpCysSerLysAlaCysLysC 1387
692 TCCTGTAGTGGCGCTGTGAGCGCGCGCGCTGCGACAGCG 646
1387 sValCysProAlaInLysLysCysaspSerProLysThrlTrpaspL 1404
645 GCAGGTATTTCGCGCTGCTCTCG...GCCGAGGTTCTTCAGGCTCA 599
1404 snserCysSerCysglInCysProLysAsnMetArgProProThrlglY 1420
598 GTTCACATTTCTTCTCTGCTGAGCGGCTCTGGCGCGCTCAAGGCC 549
1421 CysAsnAlaIglYArgThrlTrpAspAlaThrcysSerglulYcysAl 1437
548 TGGTGTCTTCATCAGCAGTGTGCGGAGGAGGTGCCCGCGGAGCAC 499
1437 aAlaThrlProLys.....CysaspSerProLysValPheaspProT 1451

```

```

498 GCACCTGAGCCGACAGAGCTACCCAGCGGCGCTGCGATCACCCGCGCC 449
1451 hTrThrcysglYcysLysCysglYasnLysProAsnLysLeuAspAlaIgl 1467
448 rCATCTTGAGGCGCTGTGTGAAGCCGCGCTGCTGCCACTGCTGGTCG 399
1468 ArgTrhTrp..... 1470
398 GGGTCTCTTGACCCCGACAGG.....AGCTGCGGCGCCGAGCT 358
1471 .....AspglulInLysCysLysMetThrcysAspLeuProAlaI 1484
357 CTCTTCCCTGGGCTAATGTCAT..... 335
1484 lerpCys.....CysHisTrglulYglInValTraspSerAsn 1497
334 .....CCAGCAGCAGAGGCT..... 317
1498 ThrcysLysCysglYcysProLysglulInThrcysLysglInlYpHese 1514
316 ...AGTTGACGCTCAGGCTGCCGAGAGCGCTGTGGCTTGCACAGCT. 272
1514 rPheSerProLysSerLysLys.....CysIleLeuG 1526
271 ....ACAGCGCGCATCTCCCGCTCCACCTGCTTACCTTACGCCCT 227
1526 lYcysasnLysaspPro..... 1532
226 GCGGACAGCAGCGGAGAGCGGCTCCAGCCCTGTGATGTGGGCGCATCC 177
1533 .....glYcysglYalalYasnLysIle...TrpCysglInlunh 1544
176 TTGCTCACATGTGTGACGCGCGGCGGCTCCCTCCTCAGTGGGCTGAG 127
1544 rCysLysCys...glucYalaserLupProAlaMetglY..... 1557
126 CCGCATAGTGGCGGCCA 110
1558 .....CysglYPro 1560

seq_name: SwissProt_39:PGBM_MOUSE
seq_documentation_block:
ID PGBM_MOUSE STANDARD; PRT; 3707 AA.
AC 005793;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE BASEMENT MEMBRANE-SPECIFIC HEPARAN SULFATE PROTEOGLYCAN CORE
DE PROTEIN PRECURSOR (HSPG) (PERLECAN) (PLC).
GN HSPG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=92078153; PubMed=1744087;
RA Noonan D.M., Fulle A., Valente P., Cai S., Horigan E., Sasaki M.,
RA Yamada Y., Hassell J.R.;
RA "The complete sequence of perlecan, a basement membrane heparan
RT sulfate proteoglycan, reveals extensive similarity with laminin A
RT chain, low density lipoprotein-receptor, and the neural cell adhesion
RT molecule."
RL J. Biol. Chem. 266:22939-22947(1991).
RN [2]
RP SEQUENCE OF 940-1601 AND 1870-2600 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=89034110; PubMed=2972708;
RA Noonan D.M., Horigan E.A., Ledbetter S.R., Vogel J.G., Sasaki M.,
RA Yamada Y., Hassell J.R.;
RA "Identification of cDNA clones encoding different domains of the

```



```

FT DISULFID 1295 1304 BY SIMILARITY.
FT DISULFID 1307 1322 BY SIMILARITY.
FT DISULFID 1363 1372 BY SIMILARITY.
FT DISULFID 1365 1379 BY SIMILARITY.
FT DISULFID 1582 1591 BY SIMILARITY.
FT DISULFID 1594 1610 BY SIMILARITY.
FT DISULFID 1613 1628 BY SIMILARITY.
FT DISULFID 1615 1638 BY SIMILARITY.
FT DISULFID 1641 1650 BY SIMILARITY.
FT DISULFID 1653 1668 BY SIMILARITY.
FT DISULFID 1792 1839 BY SIMILARITY.
FT DISULFID 1886 1932 BY SIMILARITY.
FT DISULFID 1976 2021 BY SIMILARITY.
FT DISULFID 2073 2118 BY SIMILARITY.
FT DISULFID 2170 2215 BY SIMILARITY.
FT DISULFID 2268 2313 BY SIMILARITY.
FT DISULFID 2365 2413 BY SIMILARITY.
FT DISULFID 2456 2506 BY SIMILARITY.
FT DISULFID 2554 2599 BY SIMILARITY.
FT DISULFID 2641 2686 BY SIMILARITY.
FT DISULFID 2831 2876 BY SIMILARITY.
FT CARBOHYD 65 65 O-LINKED (GLYCOSAMINOGLYCAN) (POTENTIAL).
FT CARBOHYD 71 71 O-LINKED (GLYCOSAMINOGLYCAN) (POTENTIAL).
FT CARBOHYD 76 76 O-LINKED (GLYCOSAMINOGLYCAN) (POTENTIAL).

```

alignment\_scores:  
 Quality: 293.00 Length: 836  
 Ratio: 0.823 Gaps: 43  
 Percent Similarity: 42.584 Percent Identity: 22.967

alignment\_block:

US-09-598-042a-2 x PGBM\_MOUSE

Align seg 1/1 to: PGBM\_MOUSE from: 1 to: 3707

```

112 GCGCCGCACTATGCGGCTGCAAGTGCAGAGGAGGAGCCGCCGCGCT 161
    |||:||||| |||:||||| |||:|||||
2547 GLYHISLHRLAHLRLHLSHLSYSSERLALHLCGLYASNPRLPRORPO.. 2562
162 GACCATGTGCACCAAGATGCGGCAACCATCCACAGCGGCTGGAGCGGCT 211
    |||:||||| |||:||||| |||:|||||
2563 .....,ThrlleHls.....TrpserLysL 2569
212 TCCGCGTG...CTGCGG.....CAGGCGCTGAG 237
    |||:||||| |||:||||| |||:|||||
2569 euatgalaIarProleuProTgPrlnHlsarGLleGLUGLYASNPRLHleVal 2585
238 GTGAAGCAGGTGAGCGGAGAGATCCGCGGTGACGTGCAAGGCCAC 287
    |||:||||| |||:||||| |||:|||||
2586 lIeProArGyAlaIaGLInGLInAspserGLyInTgIleCyasaAlaIaTh 2602
288 CAAGCGGTGGCGGAGCTGAGCGCTACACTACACCTGCTGCTGGATG 337
    |||:||||| |||:||||| |||:|||||
2602 rAsnSerAlaGLyHlsThrgIuaIaThr.....ValValLeuHlsV 2616
338 ACATTAACCCAGGAGAGAGAGACCTGGGCGCCGACAGCTCCTGGGGGT 387
    |||:||||| |||:||||| |||:|||||
2616 aIgluserProProTgYAlaIaThrlleIleProGLInHlsThsSer..... 2630
388 CAAGAAGACCCCGCCAGCCAGCACTGGGACAGCCGCGCTTCACAGACC 437
2630 ..... 2630
438 CTCGAAGTAGAGCGCGGATGATCGACGCGCGGTGGTAGTCTGCGTGC 487
    |||:||||| |||:||||| |||:|||||
2631 .....AlaGLInPro...GLYASNPLeuValG 2638
488 GGTCAAGTGCCTGCGGAGCGGAGCACTCCTGGCGCCGACATCACCTGGATG 537
    |||:||||| |||:||||| |||:|||||
2638 lInleuGLInCyLeuAlaIaHlsGLyThrProIleuThrlTgIlnTrpser 2654
538 AAGAGCAACCAAGCGCTTGACGCGCCGACAGGCGCGCTGAGCCAGGAAGA 587

```

```

2655 LeuValGLyGLyValLeu.....ProGLInYsaAlaValaIaArgsnGL 2669
    |||:||||| |||:||||| |||:|||||
588 GAAGTGCAGACTGAGCTGAGAACCTGGCGCGGAGGAGCAAGCGGAAT 637
    |||:||||| |||:||||| |||:|||||
2669 .....LeuValArgLeuGLInProThrlValProGLInAspserGLyArgT 2684
638 ACACSTGCGCGGTGCAAGCGGAGCGGCGCCATCCAGCCACCTGCAAG 687
    |||:||||| |||:||||| |||:|||||
2684 YArgCyGLInValaIaSerAsnArgValaIaGLySerAlaGLInaIaHlsGLIn 2700
688 GTGATGTGATCCAGCGGAGCCGCTTCAAG..... 717
    |||:||||| |||:||||| |||:|||||
2701 Val...LeuValGLInGLySerSerAsnLeuProAspThrlleIlePr 2716
718 .....CCGCTGCTACAGGCAAGCCCGGAGCAAGCGG 754
    |||:||||| |||:||||| |||:|||||
2716 oGLyGLySerThrlProThrlValaIaGLInValaIaThr...ProGLInleuGLInTha 2732
755 TGACATTCGCGGAGGAGCAAGCTCCTCCAGTGCAGAGTGCAGAGCAAGT 804
    |||:||||| |||:||||| |||:|||||
2732 rGAsnIleGLyAlaSerValaIaGLInPheHlsCyAlaValaIaProAsnGLyArg 2748
805 AAGCGCGTATCCAGTGCSTGAAG..... 828
    |||:||||| |||:||||| |||:|||||
2749 GLYThrlHlsLeuArGLTgPrlleuYsGLInGLyGLInleuProProGLIn 2765
829 .....CCGCTGGAG..... 837
    |||:||||| |||:||||| |||:|||||
2765 sSerValaIaAspGLyValaIaLeuArGLleGLInaIaSerGLInaIaAsnCyAs 2782
838 .....TACGCGCGGAGGCGCGCCGACCAACTGC 864
    |||:||||| |||:||||| |||:|||||
2782 lInGLyThrlTgYValaIaGLInaIaHlsGLyProTgPrlGLInaIaGLInaIa 2798
865 ACC..... 867
    |||:||||| |||:||||| |||:|||||
2799 ThrlaGLInleuIleValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 2815
868 .....ATCGATGTGGG...GGCGAAGATTGTGTGTGC 898
    |||:||||| |||:||||| |||:|||||
2815 gThrlSerValaIaHlsSerValaIaValaIaGLInHlsSerValaIaGLInPheGLyAl 2832
899 TGCCACAGCGGTGAC.....GTGTGTGCGCGCGCCGACAGCGC 933
    |||:||||| |||:||||| |||:|||||
2832 euAlaLeuGLInAspProGLyAspProGLInValaIaThrlTrpserGLyVal...GLy 2847
934 TCCTACCTCAATAG.....CTGCTCATPAC 959
    |||:||||| |||:||||| |||:|||||
2848 GLYHlsLeuArProGLyIleValaIaGLInserGLYThrlleIleArGLInleI 2864
960 CCGTGCSCCGCAGAGAGATGCGGAGCAATGTACATGTGCTTGGCCGACAA 1009
    |||:||||| |||:||||| |||:|||||
2864 aHlsValaIaGLInleuAlaAspAlaGLyInTgYArgCyAlaIaIaThAsn 2881
1010 CCAATGGGCTACAGCTCCGAGCGCGCTTCTCACCGTGTGCGGAGACCA 1059
    |||:||||| |||:||||| |||:|||||
2881 lalAaGLyThrlThrlGLInSerHlsValaIaLeuLeuValaIaIaIaIaIaIa 2897
1060 AAAGCGCAGAGGCAAGCTGCGGCTGCTGCTGCGGCGCACTAGCTGGC 1109
    |||:||||| |||:||||| |||:|||||
2898 GLInIleSerThrlProProGLInIleArgValaIaProIaGLySerAlaIaVala 2914
1110 GTGCGCGGTGATCGGAGCAATCCAGCGGAGCGGTGCTTATCCTGGGCA 1159
    |||:||||| |||:||||| |||:|||||
2914 lPheProCyMet..AlaSerGLYArgProThrlProAlaIaIeThrlTrpser 2930
1160 CCGTGTCTGTGCGCTTGCAGAGCCGACGAAGAAGCGGTACACCCCGCG 1209
    |||:||||| |||:||||| |||:|||||
2931 LysValaAsp...GLYAspLeuProProAspSerArgLeuGLInaIaAsnMe 2946
1210 CCGCGCGCGTCCCG.....TGCC 1226
    |||:||||| |||:||||| |||:|||||

```





41 ..... 41  
2760 CSAAGGCGAGGAGAGAGCGGGATGTGAGCGGGAGTGAAGCGGGA 2711  
42 ..... GlnProGlyGlnIsthrValGlyAlaGlyValGly 54  
2710 CAGAGCGGCGAGAGTTGGGGAGTGAAGTGGAGCGGAAAGTCCCTC 2661  
54 SerProSerSerGlnLeuTYGlnIsthrValGlnGlyGln 70  
2660 ATCAAGTATGAGCTTCAGGCGACAGTACTGCGAGCTTCAGGCGACACC 2611  
71 ValValPheThr ..... HisArgIleAsn 79  
2610 TCCCAAAAGAGCTGAGTGCCTGTGTCCAGGCGACATATGACCTG 2561  
79 uProProSerThrGly...CysGlyCysProProGlyThrGlnProPro 95  
2560 TGTGTGCATGT ..... 2550  
95 AlLeuAlaSerGlnValGlnAlaLeuArgValArgLeuGlnIleLeuGln 111  
2549 ..... GTGTCCGGAAGTGTGCCCCAG 2527  
112 GlnLeuValGlyGlyLeuGlyGlnGlyGlnGlyGlyCysGlyPro 128  
2526 CAGCATATGTGATGTGTGCGTGCCTGTATCCGCGACAG ..... 2487  
128 AserAlaGlnAla ..... GlyThrGlyGlnThrAspVal 140  
2486 ..CATATGTGCAGCTGTGTGCATGTCCAGACAGCATATGTGCACATGT 2439  
140 rGThrLeuGlySerLeuIsthrValPheAspSerLeuArgCysThrCys 156  
2438 GTGTGT...CGAGGC..AATATGTGCAGCTGT .....GTAGTGT 2404  
157 SerCysGlnProGlyThrGlyGlyProThrCysSerAspProThrAspAl 173  
2403 TGAGGAGAGCATTTATCTGTGTGTGTGTCCAGGCGACATATGTGCGTGTG 2354  
173 AgGlnIleProProSerSerProProSerAlaSerGlySerCysProAspA 190  
2353 TGTGTGTCCGAGACATATCTGTGCATGCGTGTGTGTGTCCGAGAGA 2304  
190 sPcysAsnAspArgGlnGlyArg .....CysValArgGlyArg 201  
2303 TGTGTGCGTGTGTGTGTGACATAGACAGCATATCCGGGTGTGTCCAGGCG 2254  
202 Cys .....ValCysPheProI 207  
2253 CATATGTGCGCTGTGCACAGTGTGTGTGAGAGTGTGTCCAGGCGACAT 2204  
207 yThrThrGlyPro ..... 211  
2203 ATGTGCATGTGTGTGCGTGTCCAGACAGCATATCTGTGCACAGCTGTGTG 2154  
212 .....SerCysGlyThrProSer ..... 217  
2153 TGTGTGTGTCCAGGCAATATCCGTGCATGTGTGTGTCCAGGCGACATTTAT 2104  
218 .....CysProGlyAsp .....CysGlnGlyArgGlyArg 227  
2103 CTGTGTGTGTCCAGGCGACATATCTGTGCACAGTGTGTGTGTCCGAGTACATAT 2054  
227 gCysVal .....GlnGlyValCys .....ValCysArgAlaGlyPheS 240  
2053 CTTCAGCTGTGTGCGTGTCCAGACAGCATATCCGTGTGTGTGTGTGTGTCCA 2004  
240 SerGlyProAspCysSerGlnArgSer .....CysPro 250  
2003 GCGAGCGACATTTGCCGATGGGTGTGCGTGTGTATGTTCAGGCGACATGTC 1954  
251 ArgGlyCysSerGlnArgGly .....ArgCysGlnGlyGly ..... 262  
1953 CTGTGTGTGTGCAGCATGTCTGTGTGCGTGTGCGTGCATTTGGGC ..... 1909  
263 .....ArgCysValCysAspProGly...TYThrGlyAspAs 274  
1908 ...AGCTTATCTGTGTGTCCAGGCGGCGATATCTGTGCATGTGCGTGTG 1864  
274 pCysGlyMetAlaSerSerCysProAlaGlyCysSerGlnArgGlyArgCysG 291  
1863 .....TGCGTACGTGTGCGT .....TCAGGAGGAGCGAGT... 1834  
291 LuAsnGlyArgCysValCysAsnProGlyTYThrGlyGlnAspCysGly 307  
1833 GCGCGCATGTGTGTGCATATATGCATTCAG...GTAGTGTGTG...TC 1790  
308 ValArgSerCysProAlaGlyCysSerGlnArgGlyArgCysGlyAspGly 324  
1789 TGTGTGTGTGT...CGAGGCGCTATGCTCTACACAGACAGTGCCTGG 1743  
324 yArgCysValCysAspProGly .....TYThrGlyGlnA 336  
1742 GTGCGTGCATTTCCCTCCCGGCGATGGCGCCCGCTGTGTGTGCAGCTC 1693  
336 sPcysGlyThrArgSerCysProThr...AspCys ..... 346  
1692 CGTCCCTCCATCCTCCAGTATGCTGTGTGTGCGCGGCCCGCTGTGCTC 1643  
347 .....GlyGlnGlyGlyArgCysValAspGlyArgCysVal ..... 358  
1642 TGCAGATACGCTGCCGTGTACAGTATATGATGTGTGTGTGTGACCTT 1593  
359 .....CysThrProGlyTYThrGlyGlnAspCysSerThrArgThrC 373  
1592 GCGCCCGACAGTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGGA 1543  
373 ySPPro .....ArgAspCysArgGlyArgGlyArgCysGlnAspGly 386  
1542 ...TGTGTGTGTAGAGTTGGGATAC .....AACTTAGGCGCCAGCA 1505  
387 GlnCysIleCys..AspThrGlyTYSerGlyAspAspCysGlyValArg 402  
1504 ACTGGCGCTGGGCGCCAGTAAG ..... 1484  
403 SerCysProGlyAspCysGlnArgGlyArgCysGlnAspGlyArgCys 419  
1483 ...TGCTGG...GGGCGTCCCGGAGACCCATGCTCTCCACACAGCGGC 1441  
419 sValCysThrProGlyTYThrGlyThrAspCysGlySerArgAlaCysP 436  
1440 CG .....CAATACAGTGTGTGTGTGT 1418  
436 roArgAspCysArgGlyArgGlyArgCysGlnAsnGlyValCysValCys 452  
1417 GGA ..... 1415  
453 AsnAlaGlyTYSerGlyGlnAspCysGlyValArgSerCysProGlyAs 469  
1414 .....TGCTGTGTAGAGTTT 1399  
469 pCysArgGlyArgGlyArgCysGlnSerGlyArgCysMetCys .....T 484  
1398 GGGGTACAACTTAGGCGGCAAGTGGGCTT ..... 1368  
484 rProGlyTYThrGlyArgAspCysGlyThrArgAlaCysProGlyAsp 500  
1367 .....GGGCCCACTAAGTGTGTG...GG 1348  
501 CysArgGlyArgGlyArgCysValAspGlyArgCysValCysAsnProGly 517  
1347 GGTCCCGGAGACCATGCTCTCTACACAGCCCGCACACACAGG ..... 1305  
517 yPheThrGlyGlnAspCysGlySerArgArg...CysProGlyAspCysA 533

1304 .....CCAGCGCTGAGGGCGGCGCAACGAGGAGG 1275  
533 rgglYhIsIglYleucysgluAspGlyValAlcysAspAlaIglYTr 549  
1274 TCCTGTCTCCGCTGCGGTCGGGCGCCGCCGCGGCGGCGGCCAGG 1225  
550 SerGlygluAspSerThrArSerSerProGlyGly...CysArgI 565  
1224 CAGGGG.....G 1217  
565 yArGlyglYncysleuAspGlyArGcysValcysgluAspGlyTrSerG 582  
1216 GGGCAGCGCGCGGGGTGCAC..... 1197  
582 YgluAspCysglYValArgGlnCysProAsnAspCysSerGlnHIsGly 598  
1196 .....GCTTCTTCTGGGCGCTGCAAGCCACAGCAGCAGGCT 1159  
599 ValCysgluAspGlyValcysIleCysTrpGluGlyTrValSerGluAs 615  
1158 GCCCAGGATGAGACAGCGCGGCTGGGATGCCGATGACCACGCGGCAG 1109  
615 pCysSerTrIeArGTrHcysProSerAsnCys.....HisG 627  
1108 GCAGGCTAGTGGCGGAGGAGGAGGAGCCACAGTGCCTGGCGGCTTTr 1059  
627 YArGlyArGcysgluGluGlyArGcysleucysAspProGly..... 641  
1058 GGGTCTGCAGCAGCAGGTGAGGAGGCGCTGCGGAGCTGTAGCCCATGGT 1009  
642 TrpThrGlyProThrCys.....AlaTrpArG.MecCysProAlaAspC 656  
1008 GTTGGGCGCAAGCAGATGTACA.....TGCCCGCATCG 974  
656 yArGlyArGlyArGcysValGlnGlyValcysleucysHisValGly 672  
973 CCTGGCGGACAGGTGATGACAGCTTATGAGTAGAGCCGTCGGGC 924  
673 TrpGlyGlyGlu..... 676  
923 CGGACACACAGTCACCCGTGGGACACCAAACTTGGCGGCCAC 874  
676 ..... 676  
873 ATCGATGTGAGTGTGGGGCGCTCGGGCGCTACTCCACGCGCTTCA 824  
677 .....AspCysGly.....GlnGluGluProProAlaSer 666  
823 GCCACTGATCACCGGCTTCACGTCGTGGGACCTTGACATGGAAGAC 774  
687 Ala.....CysProGlyGly 692  
773 GTGGTCCCCCGAATCCAGCTGTCACGGGTGCGCTGAG 724  
692 s...GlyProAlaGlu.....leucysArGlaGlyGlnCysValCysV 706  
723 CAGGCGCTGGAACGGGTCCGTGATCACATCCACTGTAGTGAGCG 674  
706 alGluGlyPheArGlyProAspCysAlaIleGlnThrCys..... 719  
673 TGATGGCGCGCGGCTGCACACGCGGAG..... 641  
720 .....ProGlyAspCysArGlyArGlyGluCysHisAspG 732  
640 .....TGATATTGCGCGTCTCCGCGCGAGGTCTTCAAGCTCAGTG 556  
732 ySerCysValCysAspGlyTrAlaGlygluAspCysGlyGluAla 749  
595 .....TCACCTTCTTCTGGGCTGAGCGGCTGAGCGGCGGTCA 554  
749 rgValProSerSerAlaSerAlaTrAspGlnArGlyleuAla..... 763  
553 AGGCTGTGTCTCTTCATCCACGTATGTGGGCGGAGGTCGCCGCTG 504

764 ..... 765  
503 GCCACGCACTTAGCGCGCACGAGCTACCCAGCGCGCGCGATCACCG 454  
766 .....GlnGluTrpGlnValTrValAlaGalaLeu 776  
453 GCGCCTCATCTTGAAGGCTGTGAAGCGGTCGTGCCACTGTGGC 404  
776 rgGlyThrSerTrp.GlyleuProAlaSerLysThrIleThrMetI 792  
403 TGGCGG.....GTCTTGTGACCCCA..... 379  
792 eaSpGlyProGlnAspLeuArgValAlaValThrProThrLeuG 809  
378 .....GGACTGTGGCGCCGCGCT..... 358  
809 luLeuGlyTrpLeuArGProGlnAlaGluValAspArGpHeValIser 825  
358 ..... 358  
826 TrpValSerAlaGlyAsnGlnArgValArgLeuGluValProProGlu 842  
357 .....CTCTTCCCTGGGCTAAATGTCATCCAGC 330  
842 aaSpGlyThrLeuLeuThrAspLeuMetProGlyValGlyTrValVal 858  
329 ACGACAGGCTGTAGTTCAGCTCAGCTGCGGAGCGCTGGGCGCTT 280  
859 .....ThrValThrAlaGluArgGlyArGlaValAlaSerTrpPro 871  
279 GCACAGTACAGCGCGCATC..... 259  
872 AlaSerValArgAlaAsnThrGluGluArgGluGluSerProProAr 888  
259 ..... 259  
888 pProSerLeuSerGlnProProArGArProTrpGlyAsnLeuThrAlaG 905  
258 .....CTCCGCTCCACCTG... 244  
905 luLeuSerArGpHeArGlyThrValGlnAspLeuGluArGHisLeuArG 921  
243 CTTCACCTTCAGCCCTCGCGCAGCAC..... 217  
922 AlaHisGlyTrpProLeuArGlaAlaAsnGlnThrTrpThrSerValAlaAr 938  
216 .....GCGGAGCGGCTCCAGCCGCTGTGATG 189  
938 gHisIleHisGluTrpLeuGlnArgGlnValLeuGlySerSerAlaAspG 955  
188 GTGGCGCATTCCT.....GTCACATGTGTCAGCGGCGGCGGTC 148  
955 YAlaLeuLeuValSerLeuAspGlyLeuArGlyGlnPheGlnValArgVal 971  
147 CCGCTCCACTGGGCACTGACGCGCATAGTCGGCGCCACGCGGCACT 98  
972 ValLeuArGTrpArGProGlnProProAlaGluGlyProGlyGlyLeu 988  
97 GCGGAGCAATCGTCGACCGG..... 76  
988 uThrValProGlyThrTrpArGThrValSerLeuProAspLeuArGProG 1005  
75 ..GAACGACATCAGTGGAT...GAAGACAATCCGCGCGCTGCGCG 31  
1005 YThrThrTrpHisValGluValHisGlyValArgAlaGlyGlnThrSer 1021  
30 CCGGACATGCG.....ACGTGGGCGCA 7  
1022 LysSerTrpAlaPheIleThrThrThrGlyPro 1032

seq\_name: SwissProt\_39:FR2\_DROME





```

126 GGTGAGTGGCCAGTGGAGGGGAGCCG...CCGGCGTGAACCATGTGA 172
    |||
    |||
259 rllgclucylleuleugluaspseralaleucluprolustlethtrpL 276
    |||
    |||
173 CCAAGGATGGCCG...ACCATCCAGCGGCTGGAGCGCTTCGCG 216
    |||
    |||
276 euHlsylslnlaaspasnlleasprleuclnlarleuarg 292
    |||
    |||
217 GTGCTG... 222
    |||
    |||
293 gluglnserglleuprovalaspvalthargleulletharargmetas 309
    |||
    |||
223 ...CCGAGGGCTGAAGTGAAGAGGCTGGAGCGGAGATGCGCGG 268
    |||
    |||
309 pgluproglnvalleuargleuglyasnvalleumetgluasprglyLT 326
    |||
    |||
269 TGTAGTGTGCAAGCGCCACAGCGCTTCGCG...AGCCTGAAGCTCAAC 315
    |||
    |||
326 rpyrllecylslealaglualsnglnvalglYargrhvalalaaser 342
    |||
    |||
316 TAC...ACCCCTGCTGCTGTGATGACAT 341
    |||
    |||
343 tyValaspLeuTyserProseraspTrhtrhValargThrTh 359
    |||
    |||
342 TAGCCCGAAGAGAGAGCTGGGCGCCAGACCTCTCTGGGGGTCAAG 391
    |||
    |||
359 rtrhtrhtrhValAlaserProleProthralaserThnglygluaspa 376
    |||
    |||
392 AGGAC...CCGCGCACCGCAGCGTGGCA 417
    |||
    |||
376 snaspraspaspValgluasProalaalaaspralaserlyglvalgly 392
    |||
    |||
418 CGAGCGGCTGCACAGCCCTCCAGATGAGGCGCGGCTGATGCACAG 467
    |||
    |||
393 ProProvalPhearglyscleuleuylsargleuclnHiserleuser... 408
    |||
    |||
468 GCCGTGGTGGTCCGTGGCGGCTCAAGTCCGTGGCAGCGGCGACCTTC 517
    |||
    |||
409 ...GlyAsnthrValasnleuclacysProvalTyrgly...L 421
    |||
    |||
518 GCCCGCATCAGCTGATGAGAGCAGCAGCCTTGACCGCCCGCAGAG 567
    |||
    |||
421 ysAlasnllethtrpThrLysaspLysProleuasna9...Glu 436
    |||
    |||
568 GCCGTGAGGCCAGGAAGAGAGTGAAGTGAAGCTGAGAGACCTGGC 617
    |||
    |||
437 leuclvallyValgluLysasntrpThrleuargPhevalglunlarth 453
    |||
    |||
618 GCCGAGAGAGCGGCAATACACTGCCGTGGTGCAGACCGCGCGGCG 667
    |||
    |||
453 rsergluasprserglyleuTyasnclLysValcysasnlatrpglyC 470
    |||
    |||
668 CCATCAAGCGCCTACAGAGTGTGATCCAGCGGAGCGGCTGCCAAG 717
    |||
    |||
470 ysileglnpheasprPheSerValglInlleasnasprgtrhtraserla 486
    |||
    |||
718 CCGGTGCTCAGAGCAGCAGCCCGTGAACAGCAGCGTGGACTTGGGGG 767
    |||
    |||
487 Prolelle...ValValProglInasnglnThrValLysValasngl 501
    |||
    |||
768 GACCACGCTCTTCAGTGCAGAGTGCAGCAGCAGCGTGAACCGGTGATCC 817
    |||
    |||
501 yserleuValmetLyscysThrValTyserAsprleuclnHisrProthValS 518
    |||
    |||
818 AGTGGCTGAAGCGGCTGAGTACGGCGCGAGGGCGCCAGCAACTCCACC 867
    |||
    |||
518 ertrpLysargValValleuLysasnAlaserleuasprglyleuclnser 534
    |||
    |||
868 ATCGATGTGGCGCGCAGAGTTGTGTGTGCTGCCACGGGTGAGCTGTG 917
    |||
    |||
535 ValgluileglnasleuasnpheThrValThrasnasprerValVal... 550
    |||
    |||
918 GTGGCGGCCGAGCGCTCTACTCAATTAAGTCTCATACCCGTCGCC 967
    |||
    |||

```

```

551 .....leuThrleuargsnValT 557
    |||
    |||
968 GCCAGAGATGGCGGCGATGTACATCTGCTTGGCGCCACACCATGTGGCG 1017
    |||
    |||
557 hrPheaspgluclueLyrtrpLyscysleuAlaserSerleuLysleu 573
    |||
    |||
1018 TACAGCTTCGCGAGCGCTTCCTCACCGTGTGCAGACCGCAAAACCGCG 1067
    |||
    |||
574 ArgserAsnserSerValTyrlleuargValVal...SerProleuProPr 589
    |||
    |||
1068 A.....GGCCACACTGTGGCTCTCTGCTCT 1093
    |||
    |||
589 oleucluleLyrAlaleuLeuclnHisAlasprleuclnLyrPheThrleua 606
    |||
    |||
1094 CGGCGACTAGCTGCGCGTGGCGCGCGTGCATCGGATCCAGCGCGGCT 1143
    |||
    |||
606 lalAllethrlleval.....Ala 612
    |||
    |||
1144 GTTTCATCTTGGCGACCTGCTCTGCTGTGC..... 1179
    |||
    |||
613 LeuPheleuLysSerAlaphellethrhPheMetleuargLeuar 629
    |||
    |||
1180 .....CAGGCCCAGA 1189
    |||
    |||
629 gargluLysleuLeuLysleuarglleuLuthrValHisglntrpThL 646
    |||
    |||
1190 AGAGCCGTGCACCCCGCGCTGCGCCCTCCCTGCTGGCGCGCGCG 1239
    |||
    |||
646 yslsVallelleLyrtrpProgly.....GlyluclnLys 658
    |||
    |||
1240 CCGGGAGCGCGCGCGAGCGCAGCGAGACAGACCTTCCCTGCTG... 1287
    |||
    |||
659 serglyCysSer.....serglyasprleuclnMetProvalillear 672
    |||
    |||
1287 ..... 1287
    |||
    |||
672 glleuLysglinargrThrThralserThrThrglyThrglyThra 689
    |||
    |||
1287 ..... 1287
    |||
    |||
689 sprProalaglInglyPheasngluTyrgluPheProleuasprserasntrp 705
    |||
    |||
1288 .....GCCGCTCAGCGCTGCGCTGCTGGCG... 1317
    |||
    |||
706 glulleProarglnleuLysleuLysleuLysleuLysleuLysleu 722
    |||
    |||
1318 .....CTGTGAGAGAGCATGGTCTCCGCGAGCGCGCC 1351
    |||
    |||
722 apheglYargValValMetAlaglualaglueLyrleuProargserPro 739
    |||
    |||
1352 AGCACTTACTGGGCCCGAGCGCGCTGCTGCGCTAAGTTGTACCCCAAA 1401
    |||
    |||
739 ln.....LeuAlagluthrLleValAlaValys 748
    |||
    |||
1402 CTCTACAGACATCCACACACACACACTGATGTGCGCGCGCTGCTG 1451
    |||
    |||
749 MetVallysglnleuclnHisThrasprasp.....MetAlaserleuVa 763
    |||
    |||
1452 GAGGAGCATG 1461
    |||
    |||
763 lArggluMet 766
    |||
    |||
seq_name: swissprot_39:DCG_HUMAN
seq_documentation_block:
ID DCG_HUMAN STANDARD: PRT: 1447 AA.
AC P43146.
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TUMOR SUPPRESSOR PROTEIN DCC PRECURSOR (COLORECTAL CANCER SUPPRESSOR).
GN DCC.

```

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID:9606;  
RN [1]  
RP SEQUENCE FROM N.A.A.  
RX MEDLINE=95011532; PubMed=7926722;  
RA Hedrick L., Cho K.R., Fearon E.R., Wu T.-C., Kinzler K.W.,  
RT Vogelstein B. ;  
RT "The DCC gene product in cellular differentiation and colorectal  
RT tumorigenesis." ;  
RL Genes Dev. 8:1174-1183(1994).  
RN [2]  
RP SEQUENCE OF 1-750 FROM N.A.  
RX MEDLINE=90100559; PubMed=2294591;  
RA Fearon E.R., Cho K.R., Nigro J.M., Kern S.E., Simons J.W.,  
RA Ruppert J.M., Hamilton S.R., Preisinger A.C., Thomas G., Kinzler K.W.,  
RA Vogelstein B. ;  
RT "Identification of a chromosome 18q gene that is altered in  
RT colorectal cancers." ;  
RL Science 247:49-56(1990).  
RN [3]  
RP SEQUENCE OF 107-472 FROM N.A. (SCRAMBLED EXONS).  
RX MEDLINE=91121517; PubMed=1991322;  
RA Nigro J.M., Cho K.R., Fearon E.R., Kern S.E., Ruppert J.M.,  
RA Oliner J.D., Kinzler K.W., Vogelstein B. ;  
RT "Scrambled exons." ;  
RL Cell 64:607-613(1991).  
RN [4]  
RP GENE STRUCTURE AND VARIANTS CARCINOMA HIS-1375.  
RX MEDLINE=94245241; PubMed=8188295;  
RA Cho K.R., Oliner J.D., Simons J.W., Hedrick L., Fearon E.R.,  
RA Preisinger A.C., Hedge P., Silverman G.A., Vogelstein B. ;  
RT "The DCC gene: structural analysis and mutations in colorectal  
RT carcinomas." ;  
RL Genomics 19:525-531(1994).  
RN [5]  
RP VARIANT CARCINOMA THR-168, AND VARIANT GLY-201.  
RX MEDLINE=94243823; PubMed=8187090;  
RA Miyake S., Nagai K., Yoshino K., Ota M., Endo M., Yusa Y. ;  
RT "Point mutations and allelic deletion of tumor suppressor gene DCC in  
RT human esophageal squamous cell carcinomas and their relation to  
RT metastasis." ;  
RL Cancer Res. 54:3007-3010(1994).  
RN [6]  
RP FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -1- TISSUE SPECIFICITY: FOUND IN AXONS OF THE CENTRAL AND PERIPHERAL  
CC NERVOUS SYSTEM AND IN DIFFERENTIATED CELL TYPES OF THE INTESTINE.  
CC -1- DISEASE: COLORECTAL TUMORS THAT LOST THEIR CAPACITY TO  
CC DIFFERENTIATE INTO MUCUS PRODUCING CELLS UNIFORMLY LACK DCC  
CC EXPRESSION. INACTIVATION OF DCC DUE TO ALLELIC DELETION AND/OR  
CC POINT MUTATIONS MAY CAUSE BOTH LYMPHATIC AND HEMATOGENOUS  
CC METASTASIS OF COLORECTAL SQUAMOUS CELL CARCINOMAS.  
CC -1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
CC -1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation --  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/announce/  
CC or\\_send\\_an\\_email\\_to\\_license@isb-sib.ch](http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch)).  
CC  
CC EMBL: X76132; CAA53735.1; -  
DR EMBL: M32292; AAA35751.1; -  
DR EMBL: M32286; AAA52174.1; -  
DR EMBL: M32288; AAA52175.1; ALT\_SEQ.  
DR EMBL: M32290; AAA52176.1; -  
DR EMBL: M63696; AAA52177.1; -  
DR EMBL: M63700; AAA52178.1; -  
DR EMBL: M63702; AAA52179.1; -  
DR EMBL: M63718; AAA52180.1; -

[illegible]

```

alignment_scores:
  Quality: 292.50      Length: 562
  Ratio: 1.194      Gaps: 22
  Percent Similarity: 43.354      Percent Identity: 23.665

alignment_block:
  US-09-598-042A-2 x DCC_HUMAN ..

Align seg 1/1 to: DCC_HUMAN from: 1 to: 1447

19  CGCATGCTCCGGCCCGCCATGCGCCGGGAGTGTCTTATCCAGCTGAT 68
    |||
    ::|||::|||::|||
130  ArgThrAlaIysValAlaValAlaIaIyProIeuAArgPheIeu.....143
69  GTGCTTCCGGGCTGCAGGATTCGTGCGCAGGTGGCGCGGTGGCGCGCA 118
    ::|||
144  .....SerGlnIhrnGluSerValIhrAlaPheMetGlyAspT 156
    ::|||
119  CTAATGCGGCTGCACTGCCCACTGAGAGGGGAGCCGCGCGCTGACCAATG 168
    |||
    ::|||::|||::|||
156  hrValIeuLeuIlyScysIuValIIdGlyIuIProMetProIhrIleIuS 172
    |||
169  TGGACCAACGATGGCCGC.....ACCATCCACAGCGGTGGAGCGCGCT 212
    |||

```



RP REVISIONS, SEQUENCE FROM N.A.  
RX MEDLINE-97294599; Pubmed-9150355;  
RA Gallehan D., Callahan R.;  
RT "The mouse mammary tumor associated gene INT3 is a unique member of  
the NOTCH gene family (NOTCH4).";  
RL Oncogene 14:1883-1890(1997).  
[3]  
RN SEQUENCE FROM N.A.  
RP  
RC TISSUE-Lung, and Testis;  
RA MEDLINE-96281668; Pubmed-8661805;  
RX Uytendaele H., Marazzi G., Wu G., Yan Q., Sassoon D., Kitajewski J.;  
RT "Notch4/Int-3, a mammary proto-oncogene, is an endothelial  
cell-specific mammalian Notch gene.";  
RL Development 122:2251-2259(1996).  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -1- DISEASE: ACTIVATED INT-3 TRANSFORMS MAMMARY EPITHELIAL CELLS.  
CC -1- SIMILARITY: CONTAINS 29 EGF-LIKE DOMAINS.  
CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.  
CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; M80456; AAB38377.1; -  
CC EMBL; U43691; AAC52630.1; -  
CC PIR: A38072; TYMVT3.  
DR HSSP; P00740; IIXA.  
DR MGD; MG1:107471; Notch4.  
DR InterPro; IPR000152; -  
DR InterPro; IPR000561; -  
DR InterPro; IPR000800; -  
DR InterPro; IPR001438; -  
DR InterPro; IPR001881; -  
DR InterPro; IPR002110; -  
DR Pfam; PF00008; EGF; 27.  
DR Pfam; PF00023; ank; 6.  
DR Pfam; PF00066; notch; 2.  
DR PRINTS; PR00010; EGFBL00D.  
DR PROSITE; PSS0088; ANK\_REPEAT; 5.  
DR PROSITE; PSS0297; ANK\_REP\_REGION; 1.  
DR PROSITE; PSS0010; ASX\_HYDROXYL; 11.  
DR PROSITE; PSS0022; EGF\_1; 28.  
DR PROSITE; PSS0186; EGF\_2; 21.  
DR PROSITE; PSS0187; EGF\_CA\_9.  
KW Differentiation; Neurogenesis; Repeat; EGF-like domain; Transmembrane;  
KW Glycoprotein; Proto-oncogene; ANK repeat; Signal.  
FT SIGNAL 1 20  
FT CHAIN 21 1964  
FT DOMAIN 21 1443  
FT TRANSMEM 1444 1464  
FT DOMAIN 1465 1964  
FT DOMAIN 21 60  
FT DOMAIN 61 112  
FT DOMAIN 115 152  
FT DOMAIN 153 189  
FT DOMAIN 191 229  
FT DOMAIN 231 271  
FT DOMAIN 273 309  
FT DOMAIN 311 350  
FT DOMAIN 352 388  
FT DOMAIN 389 427  
FT DOMAIN 429 470  
FT DOMAIN 472 508  
FT DOMAIN 510 546  
FT DOMAIN 548 584  
FT DOMAIN 586 622  
FT DOMAIN 622 656  
FT DOMAIN 658 686

FT DOMAIN 688 724 EGF-LIKE 18.  
FT DOMAIN 726 762 EGF-LIKE 19.  
FT DOMAIN 764 800 EGF-LIKE 20.  
FT DOMAIN 803 839 EGF-LIKE 21.  
FT DOMAIN 841 877 EGF-LIKE 22.  
FT DOMAIN 878 924 EGF-LIKE 23.  
FT DOMAIN 926 962 EGF-LIKE 24.  
FT DOMAIN 964 1000 EGF-LIKE 25.  
FT DOMAIN 1002 1040 EGF-LIKE 26.  
FT DOMAIN 1042 1081 EGF-LIKE 27.  
FT DOMAIN 1083 1122 EGF-LIKE 28.  
FT DOMAIN 1126 1167 EGF-LIKE 29.  
FT REPEAT 1168 1208 LIN/NOTCH 1.  
FT REPEAT 1209 1242 LIN/NOTCH 2.  
FT REPEAT 1243 1282 LIN/NOTCH 3.  
FT REPEAT 1268 1657 ANK 1.  
FT REPEAT 1661 1691 ANK 2.  
FT REPEAT 1695 1724 ANK 3.  
FT REPEAT 1728 1757 ANK 4.  
FT REPEAT 1761 1790 ANK 5.  
FT DISULFID 25 38 BY SIMILARITY.  
FT DISULFID 32 48 BY SIMILARITY.  
FT DISULFID 50 59 BY SIMILARITY.  
FT DISULFID 65 77 BY SIMILARITY.  
FT DISULFID 71 100 BY SIMILARITY.  
FT DISULFID 102 111 BY SIMILARITY.  
FT DISULFID 119 130 BY SIMILARITY.  
FT DISULFID 124 140 BY SIMILARITY.  
FT DISULFID 142 151 BY SIMILARITY.  
FT DISULFID 157 168 BY SIMILARITY.  
FT DISULFID 235 246 BY SIMILARITY.  
FT DISULFID 240 259 BY SIMILARITY.  
FT DISULFID 261 270 BY SIMILARITY.  
FT DISULFID 277 288 BY SIMILARITY.  
FT DISULFID 282 297 BY SIMILARITY.  
FT DISULFID 299 308 BY SIMILARITY.  
FT DISULFID 315 329 BY SIMILARITY.  
FT DISULFID 323 338 BY SIMILARITY.  
FT DISULFID 340 349 BY SIMILARITY.  
FT DISULFID 356 367 BY SIMILARITY.  
FT DISULFID 361 376 BY SIMILARITY.  
FT DISULFID 378 387 BY SIMILARITY.  
FT DISULFID 393 404 BY SIMILARITY.  
FT DISULFID 398 415 BY SIMILARITY.  
FT DISULFID 417 426 BY SIMILARITY.  
FT DISULFID 433 449 BY SIMILARITY.  
FT DISULFID 443 458 BY SIMILARITY.  
FT DISULFID 460 469 BY SIMILARITY.  
FT DISULFID 476 487 BY SIMILARITY.  
FT DISULFID 481 496 BY SIMILARITY.  
FT DISULFID 498 507 BY SIMILARITY.  
FT DISULFID 514 525 BY SIMILARITY.  
FT DISULFID 519 534 BY SIMILARITY.  
FT DISULFID 536 545 BY SIMILARITY.  
FT DISULFID 552 563 BY SIMILARITY.  
FT DISULFID 557 572 BY SIMILARITY.  
FT DISULFID 574 583 BY SIMILARITY.  
FT DISULFID 590 601 BY SIMILARITY.  
FT DISULFID 595 610 BY SIMILARITY.  
FT DISULFID 612 621 BY SIMILARITY.  
FT DISULFID 626 637 BY SIMILARITY.  
FT DISULFID 631 646 BY SIMILARITY.  
FT DISULFID 648 655 BY SIMILARITY.  
FT DISULFID 662 669 BY SIMILARITY.  
FT DISULFID 664 674 BY SIMILARITY.  
FT DISULFID 676 685 BY SIMILARITY.  
FT DISULFID 692 703 BY SIMILARITY.  
FT DISULFID 697 712 BY SIMILARITY.

```
FT DISULFID 714 723 BY SIMILARITY.
FT DISULFID 730 741 BY SIMILARITY.
FT DISULFID 735 750 BY SIMILARITY.
FT DISULFID 752 761 BY SIMILARITY.
FT DISULFID 768 779 BY SIMILARITY.
FT DISULFID 773 788 BY SIMILARITY.
FT DISULFID 790 799 BY SIMILARITY.
FT DISULFID 807 818 BY SIMILARITY.
FT DISULFID 812 827 BY SIMILARITY.
FT DISULFID 829 838 BY SIMILARITY.
FT DISULFID 845 856 BY SIMILARITY.
FT DISULFID 850 865 BY SIMILARITY.
FT DISULFID 867 876 BY SIMILARITY.
FT DISULFID 882 903 BY SIMILARITY.
FT DISULFID 897 912 BY SIMILARITY.
FT DISULFID 914 923 BY SIMILARITY.
FT DISULFID 930 941 BY SIMILARITY.
FT DISULFID 935 950 BY SIMILARITY.
FT DISULFID 952 961 BY SIMILARITY.
FT DISULFID 968 979 BY SIMILARITY.
FT DISULFID 973 988 BY SIMILARITY.
FT DISULFID 990 999 BY SIMILARITY.
FT DISULFID 1006 1019 BY SIMILARITY.
FT DISULFID 1011 1028 BY SIMILARITY.
FT DISULFID 1030 1039 BY SIMILARITY.
FT DISULFID 1046 1057 BY SIMILARITY.
FT DISULFID 1051 1069 BY SIMILARITY.
FT DISULFID 1071 1080 BY SIMILARITY.
FT DISULFID 1087 1098 BY SIMILARITY.
FT DISULFID 1092 1110 BY SIMILARITY.
FT DISULFID 1112 1121 BY SIMILARITY.
FT DISULFID 1130 1142 BY SIMILARITY.
FT DISULFID 1136 1155 BY SIMILARITY.
FT DISULFID 1157 1166 BY SIMILARITY.
```

alignment\_scores:  
Quality: 291.00 Length: 1170  
Ratio: 0.658 Gaps: 74  
Percent Similarity: 37.778 Percent Identity: 23.419

alignment\_block:

US-09-598-042a-2/rev x NTC4\_MOUSE

Align seg 1/1 to: NTC4\_MOUSE from: 1 to: 1964

```
2690 GGAAGTACGGTGGAGCAGGAGGAAATCCATCATGAGCTGACG 2641
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
185 GLYHISrhcgsluargaspriLeasnGlcysrPheLeuclurPro..GI 200
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2640 CACACGACTACGAGCTTCACGGCAGACCCCTCCAAAGCAGCTGAGT 2591
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
200 yProCysrProGlnGlyThrSerCysHIsanThrLeuGlySerTyrGlnC 217
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2590 GCGGNGTCCAGGACGATATCTGCAACCTGTGTGCACTGTGTGTCGGA 2541
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
217 yslLeuCySproValGlyGlnLeuGlyProGlnCySylsLeuArgLysGly 233
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2540 AGTGTGTGCCAGCAGCAGCATATCTGCATGTGTGCTGCGTGTATCCG 2491
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
234 ALAcysrProGlnGlySer.....CysLeuasnol 243
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2490 ACAGCATATCTGCAGCTGTGTGTCATGTC.....CAGACAGCAT 2453
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
243 yGly.....ThrCysGlnLeuValProGlnGlyHIsSerThrPheN 257
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2452 ATCTGACCATGTGTGTGTCCAGC..... 2427
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
257 IslLeuCySLeuCySpro...ProGlyPheThrGlyLeuAspCySgluMet 272
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2426 ..... 2410
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
273 AsnProAspAspCySValArgHIsGlnCySglInAsnGlyAlaIatHrcysle 289
```

```
2409 GAGHGTGGAGCAGCATATCTGTGTGTGTGTGTCAGGCAGC..... 2369
      : : : : ||| ||| : : : : ||| ||| : : : :
289 uAspGly..LeuAspThrTyrThrCysLeuCySproLysrIntrPlySgly 305
      : : : : ||| ||| : : : : ||| ||| : : : :
2368 .....ATATCTGCCTGC..... 2357
      ||| : : : : |||
306 TrpAspCySserGlnuAspIleAspGlnCySgluAlaArgGlyProProArg 322
      : : : : ||| ||| : : : : ||| ||| : : : :
2356 .....GTGTGTGTGTGCCAGCAGCATATCTGTGTCATGCG 2323
      : : : : ||| ||| : : : : ||| ||| : : : :
322 gCySArgAsnGlyGlyThrCysGlnAsnThrAlaGlySerPheHIsCysV 339
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
339 alCySValSerGlyTyrGlyGlyAlaGlyCySglu.....GluAsnLeu 353
      : : : : ||| ||| : : : : ||| ||| : : : :
2272 TCCGCGTGTGCCAGCAGCAGCATATCTGCGCTGTGTCAGTGTGTGGA 2223
      : : : : ||| ||| : : : : ||| ||| : : : :
354 AspAspCySAlaAlaIatHrcysAlaProGlySerThrCysIleAspAr 370
      : : : : ||| ||| : : : : ||| ||| : : : :
2222 GTGTGTGTGCCAGCAGCATATCTGCATGTGTGTGTGTCAGCAGCAT 2173
      : : : : ||| ||| : : : : ||| ||| : : : :
370 gValGlySer.....PheSerCySLeuCySproProGlyArgThrGlyL 385
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2172 ATCTGTGCACACGCTGTGTGTGTGTGTGTGTCAGCAGCATATCCGATGT 2123
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
385 euLeuCySHisLeuGluAspMetCySLeuSerGln...ProCySHIsVal 400
      : : : : ||| ||| : : : : ||| ||| : : : :
2122 TGTGTCAAGCAGCATATATCTGTGTGTGTGTGTGTGTCAGCAGCAT 2073
      : : : : ||| ||| : : : : ||| ||| : : : :
401 AsnAlaGlnCySserThrAsnProLeuThrGlySerThr..LeuCySIlleC 417
      : : : : ||| ||| : : : : ||| ||| : : : :
2072 GTGTGTCCGATATCCATATCTGCAGCTGTGTGTGTGTCAGCAGCATATCC 2023
      : : : : ||| ||| : : : : ||| ||| : : : :
417 yslInProGlyTyrSerGlySerThrCysHIsGlnAsp..... 429
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2022 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1973
      : : : : ||| ||| : : : : ||| ||| : : : :
430 .....LeuAspGlnLucy 433
      : : : : ||| ||| : : : : ||| ||| : : : :
1972 TATGTTCAGCAGCAGCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1923
      : : : : ||| ||| : : : : ||| ||| : : : :
433 sGlnMetAlaGln.....GlnGlyProSer.....ProC 443
      : : : : ||| ||| : : : : ||| ||| : : : :
1922 GCGTGCATTTGGGACGATATCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1873
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
443 yseluHIsGlySerCysIleasnThrProGly..SerPheasnCySle 459
      : : : : ||| ||| : : : : ||| ||| : : : :
1872 GTGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1853
      : : : : ||| ||| : : : : ||| ||| : : : :
459 uCySLeuProGlyTyrThrGlySerArgCySgluAlaAspHisAsnGlnC 476
      : : : : ||| ||| : : : : ||| ||| : : : :
1852 GCGTTCA.....GGGAGCAGCTGT..... 1834
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
476 yslLeuSerGlnProCySHIsProGlySerThrCysLeuAspLeuAla 492
      : : : : ||| ||| : : : : ||| ||| : : : :
1833 .....GCGCGCATGTGT..... 1822
      : : : : ||| ||| : : : : ||| ||| : : : :
493 ThrPheHIsCySLeuCySProProGlyLeuGlnGlyArgLeuCySgluVa 509
      : : : : ||| ||| : : : : ||| ||| : : : :
1821 ....GTGCATCATCATCCAGTAGTGTGTGTGTGTGTGTGTGTGTGTGT 1776
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
509 lGlnuValasnGlnCySThrSerAsnProCySLeuasnGlnAlaIaCySH 526
      : : : : ||| ||| : : : : ||| ||| : : : :
1775 CAGGGGCTATGCGCTCACACACAGACTGCGTGGGTGTGGCAT..... 1732
      : : : : ||| ||| : : : : ||| ||| : : : :
526 lAspLeuLeuasnGlnPheGlnCySLeuCySLeuProGlyPheThrGly 542
      : : : : ||| ||| : : : : ||| ||| : : : :
1731 .....TCCTCTGTGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1712
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
543 ALaArgCySgluLysAspMetAspGlnCySserSerThrInProCySAlaAs 559
      : : : : ||| ||| : : : : ||| ||| : : : :
1711 GTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1662
```

```

559 nglyglYArgCysArgAspGlnProGlyAlaIlePheTyrCysGluCysLeuP 576
      |||   : : :   : : :   |||||
1661 CCGGCCCCCGCCGCTCTGCAGATACGAGTGCCTAGACATGATAGT 1612
      |||   : : :   : : :   : : :
576 roGlyPheGlyGlyProHisCysGlu.....LysGlyValAspGlu 589
      |||   : : :   : : :   : : :
1611 GGATGTGCTGGTGGACCTTGGCTTCCAGCTGAGTGTGTGTGACAGTGT 1562
      |||   : : :   : : :   : : :
590 ...CysLeuSerAspProCysProValGlyAlaSer..... 600
      |||   : : :   : : :   : : :
1561 GTGTGTGTGTGTGTGTA.....TGATGTAGAGTTTGGGGTA 1521
      |||   : : :   : : :   : : :
601 .CysLeuAspLeuProGlyAlaIlePhePheCysLeuCysAspPro..GlyPhe 616
      |||   : : :   : : :   : : :
1520 CAACCTTAGGGCCAGCACTGGG..... 1499
      |||   : : :   : : :   : : :
616 e.....ThrGlyGlnLeuCysGluValProLeuCysThrP 628
      |||   : : :   : : :   : : :
1498 .....CCTGGCCCACTAGTGTGGGG..... 1475
      |||   : : :   : : :   : : :
628 roAsnMetCysGlnProGly...GlnGlnCysGlnGlyGlnGlnHisArg 643
      |||   : : :   : : :   : : :
1474 .....GCTGGCCGGAGA 1464
      |||   : : :   : : :   : : :
644 AlaProCysLeuCysProAspGlySerProGlyCysValProAlaGlnAs 660
      |||   : : :   : : :   : : :
1463 CCCATGCTCTCTACACAGG...CGGCGCAATATACAGTGTGTGTGTGT. 1418
      |||   : : :   : : :   : : :
660 PAsnCysArgCysHisHisGlyHisCysGlnArgSerLeuCysValCysA 677
      |||   : : :   : : :   : : :
1417 .....GGATGTCTG 1409
      |||   : : :   : : :   : : :
677 spGluGlyTrpThrGlyProGluCysGluThrGlnLeuGlyCysIle 693
      |||   : : :   : : :   : : :
1408 .....TGTAAGTTTGGGATCACTTAGGGCCAGCACTGGGCT 1368
      |||   : : :   : : :   : : :
694 SerThrProCysAlaHisGlyGlyThrCysHisProGlnProSerGlyTyr 710
      |||   : : :   : : :   : : :
1367 GGGC.....CCAGTAACTGTGGGGGCTG.....CGGAGA 1336
      |||   : : :   : : :   : : :
710 PAsnCysThrCysProAlaGlyTyrMetGlyLeuThrCysSerGlnGluVal 727
      |||   : : :   : : :   : : :
1335 CCCATGCTCTCTACACAGG.....CAGCAGGCCAGCGCTGAGGG 1292
      |||   : : :   : : :   : : :
727 AlThrAlaCysHisSerGlyProCysLeuAsnGlyGlySerCysSerIle 743
      |||   : : :   : : :   : : :
1291 CGGCAACGAGGAGAGTCTGTCTCCGCTGC.....GG 1257
      |||   : : :   : : :   : : :
744 ArgPro.....GluGlyTyrSerCysThrCysLeuProSerHisThrGly 758
      |||   : : :   : : :   : : :
1256 TCGCGGGCGGCTCCCGCGGGCGGCTGCCAGGAGGAGGAGGAGGAGGAGG 1207
      |||   : : :   : : :   : : :
758 YArgHisCysGlnThrAlaValAlaPheHisCysValSerAlaSerCysLeuA 775
      |||   : : :   : : :   : : :
1206 GGGGCTGACAGGCTTCTTCTGGGCTGGCAAA.....GCC 1172
      |||   : : :   : : :   : : :
775 snGlyGlyThrCysValAlaAsnLysProGlyThrPhePheCysLeuCysAla 791
      |||   : : :   : : :   : : :
1171 ACAGA.....GC 1164
      |||   : : :   : : :   : : :
792 ThrGlyPheGlnGlyLeuHisCysGluGlyValThrAsnProSerCysAl 808
      |||   : : :   : : :   : : :
1163 AGGCTGCCAGAGATGAGACAGCGCGGCTGGATGCCATGACACACGAG 1114
      |||   : : :   : : :   : : :
808 AspSerPro.....CysArg...AsnLysA 816
      |||   : : :   : : :   : : :
1113 CCAGGCGCAGGCTAGTGGCCGAGAGAGAGAGCCACAGATGGCCCTGGCG 1064
      |||   : : :   : : :   : : :
816 LeThrCysGlnAspThrProArg..... 823
      |||   : : :   : : :   : : :
1063 GTTTTGGGTCTGGGAGCAGCGGATGAGGAGGCGCTGGAGCTGTAGGCC 1014
      |||   : : :   : : :   : : :

```

```

824 .....GlyAlaIleArgCys...LeuCysSerP 831
      |||   : : :   : : :   : : :
1013 ATGTGTGTGGCGCCAGGACAGATGTACATGCCCGCATGTCTTGGCGGCG 964
      |||   : : :   : : :   : : :
831 o.Gly.....TyrThrGly 835
      |||   : : :   : : :   : : :
963 ACGGCTGATGAGCAGCTTATTGAGTAGAGCGCGTGGGCGCCGAGCCACA 914
      |||   : : :   : : :   : : :
836 SerSerCysGlnThrLeuIleAsnLeuCysAlaIleGlyLysProCysProH 852
      |||   : : :   : : :   : : :
913 CGTCACCCGT.....GGCAGCAGCAGCAGCAACTTCTG..... 883
      |||   : : :   : : :   : : :
852 sThrAlaIleArgCysLeuGlnSerGlyProSerPheIleIleIleCysLeuG 869
      |||   : : :   : : :   : : :
882 .....GCCGCCACATGATGATGTGAGTTGTGGCGCCCTCGGC 844
      |||   : : :   : : :   : : :
869 InGlyTrpThrGlyAlaLeuCysAspPheProLeuSerCysGlnMetAla 885
      |||   : : :   : : :   : : :
843 GCCGTACTCCAGCGCTTCAGCC..... 821
      |||   : : :   : : :   : : :
886 AlaMetSerGlnGlyIleGlyIleSerGlyLeuCysGlnAsnGlyGlyLe 902
      |||   : : :   : : :   : : :
820 .....ACTGATCAGCGGCTTTCAGTGTGCGGACCT..... 788
      |||   : : :   : : :   : : :
902 uCysIleAspThrGlySerSerTyrPheCysArgCysProProGlyPheG 919
      |||   : : :   : : :   : : :
787 .....TGACGTGAGAGAGCTGTGTCGCCCGGAGATCCACCGTGC 749
      |||   : : :   : : :   : : :
919 InGlyLeuLeuCysGlnAsp...AsnMetAsnProCysGlnProAsnPro 934
      |||   : : :   : : :   : : :
748 TGTTCACGGGCTGCG.....TGCT 729
      |||   : : :   : : :   : : :
935 CysHisHisGlySerThrCysValProGlnProSerGlyTyrValCysGly 951
      |||   : : :   : : :   : : :
728 GTGAGCAGCGGCTTGAACAGG.....TCGGCTG 700
      |||   : : :   : : :   : : :
951 nCysAlaProGlyTyrGlnGlyGlnAsnCysSerIleValLeuGlnAlaC 968
      |||   : : :   : : :   : : :
699 GATCAATCCACCTTGTAGTGCGCTGATAGCGCGCGCGGCTTGACA 650
      |||   : : :   : : :   : : :
968 YsGlnSerGlnProCysHis.....Asn 975
      |||   : : :   : : :   : : :
649 CGCGGACAGGTAT.....TGCGGTCTCTCC 621
      |||   : : :   : : :   : : :
976 HisGlyThrCysThrSerArgProGlyGlyPheHisCysAlaCysProP 992
      |||   : : :   : : :   : : :
620 GAGCCAGAGTCTTCAAGCTCAGTCCACTTCTTCTTCTGGCTCAGC 571
      |||   : : :   : : :   : : :
992 oGlyPheValGlyLeuArgCysGlnGlyAspValAspGlyCysLeuAspA 1009
      |||   : : :   : : :   : : :
570 GGCCTCTGGGCGGCTCAAGGCTGTGCTCTTCA..... 535
      |||   : : :   : : :   : : :
1009 rgProCysHisProSer..GlyThrAlaIleAsnCysHisSerLeuAlaAsnAl 1025
      |||   : : :   : : :   : : :
534 .....CCAGCTGATGTGCGCGCGGAGGAGGCGCGGCTG 504
      |||   : : :   : : :   : : :
1025 aPheTyrCysGlnCysLeuProGlyHisThrGlyAlaIleArgCys..... 1039
      |||   : : :   : : :   : : :
503 GGCAGCAGCTTGAAGCGGACGAGCTACCCAGGCGCTGCATCAGCCG 454
      |||   : : :   : : :   : : :
1040 .....GlyValAlaIleMetAsnLeuCysGlnSerGlnProCysSerAsn... 1053
      |||   : : :   : : :   : : :
453 GCGGCTCATTTGAGGCTGTGTGAAGCGCGTGC..... 418
      |||   : : :   : : :   : : :
1054 .....GlyGlySerCysGlnIleThrThrGlyProProProGly 1066
      |||   : : :   : : :   : : :
417 .....TGCCCATCTGCTGGCGGCGGCTCTTGTGACCCCA 381
      |||   : : :   : : :   : : :
1066 yPheThrCysHisCysPro..LysGlyPheGlnGlyProThrCysSerHis 1082
      |||   : : :   : : :   : : :
380 GAGG.....AGCTGT..... 371
      |||   : : :   : : :   : : :
1082 sLysAlaLeuSerCysGlyIleHisHisCysHisAsnGlyGlyLeuCysL 1099
      |||   : : :   : : :   : : :

```

```

370 .....CGGGCCCCAGGCTCTCTCCCTGGCGCAATGTC..... 338
1099 euProserProlyserProGlyserProIleucysAlacysIleuSergly 1115
337 .....CATCCAGCAGCAGCAGGGGTGAGTGG 312
1116 PheGlyGlyProAspPcysLeuThrProProAlaProProGlyCysGlyIyr 1132
311 ACGCTCAGGCTGCCGAGACCGCTTGGTGGCCTTGACACAGTACA.....C 268
1132 opProserProCysLeuH1aSn...GlyThrCysThrGlyThrProGlyL 1148
267 GCCGGATCCTT.....CCGCTCCACCTGCTTCCACCTTCAGCC 230
1148 euGlyAsnProGlyPheGlyNcysThcCysProProAspserPro...Gly 1163
229 CCGCCGAGCAGCAGCAGCAGCGGCTCCAGCCGCTGTGATGTGCGGCCA 180
1164 ProArgCysGlnArg...ProGlyAlaSerGlyGlyGlyGlyGly 1179
179 TCCTTGTCACATGTCAGCGGCGGGTCCCTCCACTGGGCACTG 130
1179 yAsp.....GlyThrC 1183
129 CAGCCGATAGTCGGCGCCAGCGGCGACCTTCGACGAAATCGTCGAC 80
1183 yAspAlaGlyCys.....SerglyProGlyGlyAsp.....TPAsp 1195
79 CCGGAGACGACATCAGCTGATGAGACAAATCCCGCGCATGGCGG 32
1196 GlyGlyAspCysSerLeuGly.....ValProAspProThrIlys 1208

```

seq\_name: SwissProt\_39: MUC1\_HUMAN

seq\_documentation\_block:

```

ID MUC1_HUMAN STANDARD; PRT; 1255 AA.
AC P15941; P15942; P13931; P17626; Q14128; Q16442; Q16437; Q9V4J2;
DT 01-JAN-1990 (Rel. 13. Created)
DT 01-APR-1990 (Rel. 14. Last sequence update)
DT 01-OCT-2000 (Rel. 40. Last annotation update)
DE MUCIN 1 PRECURSOR (POLYMORPHIC EPITHELIAL MUCIN) (PEM) (PEMT)
DE (EPITHELIAL) (TUMOR-ASSOCIATED MUCIN) (CARCINOMA-ASSOCIATED MUCIN)
DE (TUMOR-ASSOCIATED EPITHELIAL MEMBRANE ANTIGEN) (EMA) (H2IAG) (PEANUT-
DE REACTIVE URINARY MUCIN) (PUM) (BREAST CARCINOMA-ASSOCIATED ANTIGEN
DE DF3).
GN MUC1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Pancreas;
RA MEDLINE=90368716; PubMed=2394722;
RA Ian M.S., Batra S.K., Qi W.-N., Metzgar R.S., Hollingsworth M.A.;
RT "Cloning and sequencing of a human pancreatic tumor mucin cDNA.";
RL J. Biol. Chem. 265:15294-15299(1990).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=90202794; PubMed=2318825;
RA Lichtenberg M.J.L., Vos H.L., Gennissen A.M.C., Hilkens J.;
RT "Epitallin, a carcinoma-associated mucin, is generated by a
RT polymorphic gene encoding splice variants with alternative amino
RT termini.";
RL J. Biol. Chem. 265:5573-5578(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE-Breast carcinoma;
RA MEDLINE=90368715; PubMed=1697589;
RA Gendler S.J., Lancaster C.A., Taylor-Papadimitriou J., Duhig T.,
RA Peat N., Burchell J., Pemberton L., Lalani E.-N., Wilson D.;
RT "Molecular cloning and expression of human tumor-associated

```

```

RT polymorphic epithelial mucin.";
RL J. Biol. Chem. 265:15286-15293(1990).
RN [4]
RP SEQUENCE FROM N.A.
RA MEDLINE=91097524; PubMed=2268309;
RA Lancaster C.A., Peat N., Duhig T., Wilson D.,
RA Taylor-Papadimitriou J., Gendler S.J.;
RT "Structure and expression of the human polymorphic epithelial mucin
RT gene: an expressed VNTR unit.";
RL Biochem. Biophys. Res. Commun. 173:1019-1029(1990).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE-Breast carcinoma;
RA MEDLINE=90276413; PubMed=2351132;
RA Wreschner D.H., Hareuveni M., Tsarfaty I., Smorodinsky N.,
RA Horov J., Zaretsky J., Kotkes P., Weiss M., Lathe R., Dion A.,
RA Keydar I.;
RT "Human epithelial tumor antigen cDNA sequences. Differential splicing
RT may generate multiple protein forms.";
RL Eur. J. Biochem. 189:463-473(1990).
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE-Breast carcinoma;
RA MEDLINE=90276414; PubMed=2112460;
RA Hareuveni M., Tsarfaty I., Zaretsky J., Kotkes P., Horov J.,
RA Zrihan S., Weiss M., Green S., Lathe R., Keydar I., Wreschner D.H.;
RT "A transcribed gene, containing a variable number of tandem repeats,
RT codes for a human epithelial tumor antigen. cDNA cloning, expression
RT of the transcribed gene and over-expression in breast cancer
RT tissue.";
RL Eur. J. Biochem. 189:475-486(1990).
RN [7]
RP SEQUENCE FROM N.A.
RA MEDLINE=91033045; PubMed=1688329;
RA Tsarfaty I., Hareuveni M., Horov J., Zaretsky J., Weiss M.,
RA Jeltsch J.M., Garner J.M., Lathe R., Keydar I., Wreschner D.H.;
RT "Isolation and characterization of an expressed hypervariable gene
RT coding for a breast-cancer-associated antigen.";
RL Gene 93:313-318(1990).
RN [8]
RP PARTIAL SEQUENCE FROM N.A.
RA MEDLINE=88330762; PubMed=3417635;
RA Gendler S.J., Taylor-Papadimitriou J., Duhig T., Rothbard J.,
RA Burchell J.;
RT "A highly immunogenic region of a human polymorphic epithelial mucin
RT expressed by carcinomas is made up of tandem repeats.";
RL J. Biol. Chem. 263:12820-12823(1988).
RN [9]
RP SEQUENCE OF 1-169 FROM N.A.
RA MEDLINE=90088473; PubMed=2597151;
RA Abe M., Siddiqui J., Kufe D.;
RT "Sequence analysis of the 5' region of the human DF3 breast
RT carcinoma-associated antigen gene.";
RL Biochem. Biophys. Res. Commun. 165:644-649(1989).
RN [10]
RP SEQUENCE OF 1-109 FROM N.A.
RC TISSUE-Thyroid;
RA MEDLINE=96183746; PubMed=8608966;
RA Weiss M., Baruch A., Keydar I., Wreschner D.H.;
RT "Preoperative diagnosis of thyroid papillary carcinoma by reverse
RT transcriptase polymerase chain reaction of the MUC1 gene.";
RL Int. J. Cancer 66:55-59(1996).
RN [11]
RP SEQUENCE OF 1-89 FROM N.A.
RC TISSUE-Lung;
RA MEDLINE=96181716; PubMed=8604237;
RA Yu C.U., Yang P.C., Shew J.Y., Hong T.M., Yang S.C., Lee Y.C.,
RA Lee J.N., Luh K.T., Wu C.W.;
RT "Mucin mRNA expression in lung adenocarcinoma cell lines and
RT tissues.";
RL Oncology 53:118-126(1996).
RN [12]
RP SEQUENCE OF 1-46 FROM N.A.

```





```

716 AGCCCTGCTCAGGACGACGACCCCTGACAGACAGGCTGACCTTGGG 765
    ::::: ::::: ||||| :::::
214 ProLanIsgLylValThrSerAlaProAspThrArgProAlaProIylse 230
    ::::: ::::: :::::
766 GGGACGACGCTCTCCAGTGCAGAGGTGCGACGACGTGAAGCCGGTAT 815
    ::::: ::::: :::::
230 rThAlaProProAlaIsgLylValThrSerAlaProAspThrArgPro 247
    ::::: ::::: :::::
816 CCGAGTGGCTGAAGCGGCTGAGAGGAGCGGCGGAGGCGGCGGCGG 853
    ::::: ||||| :::::
247 laProIyl.....SerThrAlaProProAlaIsgLylValThr 259
    :::::
854 .....GCGACACCTCCACCATCGAT 873
    :::::
260 SerAlaProAspThrArgProAlaProIylSerThrAlaProPro..... 274
    :::::
874 GTGGGCGGCGCAAGTTTGTGTGTGCTCCGACGCGGTGACGCTGTGTCGG 923
    ::::: :::::
275 .....AlaIsgLylValThrSerAlaProAspThrArgProAlaPro 289
    :::::
924 GCGCGACGCTCT.....ACCTGATTAAGCTGCTGATCA 958
    ::::: ||||| :::::
289 lylSerThrAlaProProAlaIsgLylValThrSerAlaProAspThrArg 305
    :::::
959 CCGGTGCGCGCGACGACG.....ATGCGGCGATGTAATCTGCGC.. 997
    ::::: :::::
306 ProAlaProIylSerThrAlaProProAlaIsgLylValThrSerAla 322
    :::::
998 ....TTGGGCGCAACATGGGCTACAGCTTCCGACGCGCTTC..CTCA 1042
    ::::: ||||| :::::
322 oAspThrArgProAlaProIylSerThrAlaProProAlaIsgLylVal 339
    :::::
1043 CCGTGTGCGCGACGCGCAAAACG...CGAGG.....CGACGTG 1080
    ::::: ||||| :::::
339 lylSerAlaProAspThrArgProAlaProIylSerThrAlaProProAla 355
    :::::
1081 GCGCTGCTGCTCTGCGCG...ACTAGCTGCGCTGCGCGCTGCTGCTACG 1127
    ::::: :::::
356 lIsgLylValThrSerAlaProAspThrArgProAlaProIylSerThrAl 372
    :::::
1128 CATCCGACGCGCGCTGTCTTCACTCGGACGCGCTGCTGCTGCTGCTT 1177
    :::::
372 aProProAlaIsgLylVal.....Ths 380
    :::::
1178 GCGAGCGCGCGAGAAAGCGCTGACGCGCGCGCTGCGCGCTGCGCT 1227
    :::::
380 exAlaProAspThrArgProAlaProIylSerThrAlaProProAlaIsg 396
    :::::
1228 GGGACGCGCGCGCGGAGCGCGCGCGACGCGGAGGAGACAGGACCT 1277
    :::::
397 gLylValThrSerAlaProAspThrArgProAlaProIylSerThrAla 413
    :::::
1278 TCGCTGCTGCGCGCGCTGACGCT.....GCGCGCTGCTG 1312
    ::::: ||||| :::::
413 oProAlaIsgLylValThrSerAlaProAspThrArgProAlaProIyls 430
    :::::
1313 TGGGCTGTGTGAGAGCATGGGTCTCCGACGCGCGCGCGACCTACTG 1362
    :::::
430 exThrAlaProProAlaIsgLylValThrSerAlaProAspThrArgPro 446
    :::::
1363 GGGCGACGCGCGAGTGTGCGCGCTAAGTTGTACCCAACTGACACAGA 1412
    :::::
447 AlaProIylSerThrAlaProPro..... 454
    :::::
1413 CATCCACACACACACACTGTATGCGCGCGCTGTGTAGAGCATGG 1462
    :::::
455 .....AlaIsg 457
    :::::
1463 GTCTCGGACGCGCGCGACGCTTACTGCGCGCGCGCGCGCGCTGTCGGC 1512
    :::::
457 lylValThrSerAlaProAspThrArgProAla..... 467
    :::::
1513 CTAAAGTTGACCCCAACTCTACACAGACATCCACACACACACACACA 1562
    :::::

```

```

468 .....ProIylSerThrAlaProProAlaIsgLylValThr.S 480
    :::::
1563 CACTCTGCACACACACTC...ACAGCTGGAGGCGAAGGCTCCACGACGAT 1609
    :::::
480 exAlaProAspThrArgProAlaProIylSerThrAlaProProAlaIsg 496
    :::::
1610 CCACTATCAGTCTGACGCGCGCGCTATCTGCAGAGGCGCGGCGGCGCG 1659
    :::::
497 gLylValThrSerAlaProAspThrArgProAlaProIylSerThrAla 513
    :::::
1660 GCGCGACACAGGACGACTGGAGAGTGCAGAGCGAGCTGCACAGAGCG 1709
    :::::
513 oPro..... 514
    :::::
1710 AGGGACCGCATGGCGCGAGAGAAATGGCGCGACGCCCGACGACTGTGTG 1759
    :::::
515 ....AlaIsgLylValThrSerAlaProAspThrArgPro..... 526
    :::::
1760 TGAGCGATGACCCCTGGACACACACACACACACAC..... 1798
    :::::
527 .....AlaProIylSerThrAlaProProAlaIsgLylValThrse 540
    :::::
1799 .....ACTACCTGGATGATGTATGCACACACATGCGCGCGACG 1838
    :::::
540 rAlaProAspThrArgProAlaProIylSerThrAlaProProAlaIsg 557
    :::::
1839 TGCTGCTGACGCGCGACGCTACGCGACACGCGACATGCAAGATATGCG 1888
    :::::
557 lylValThrSerAlaProAspThrArgProAlaProIylSerThrAla 573
    :::::
1889 CCGGCGCGACAGATA...AGCTGCCAAATGCGCGCGCGCGCA...G 1932
    :::::
574 ProAlaIsgLylValThrSerAlaProAspThrArgProAlaProIyls 590
    :::::
1933 AGACATGCCAGAAATGCAAGGACATGCTGCTGCAATGACACGCGACA 1982
    :::::
590 exThrAlaProProAlaIsgLylVal.....ValThrSerAla 601
    :::::
1983 CCGATGCGCGAGATGCTGCTGCGCGACACA.....CACACGCGAT 2026
    :::::
602 ProAspThrArgProAla..ProIylSerThrAlaProProAlaIsgLyl 618
    :::::
2027 ATGCTGTGTGACGCGACACAGT.....GCGATATG 2058
    :::::
618 alThrSerAlaProAspThrArgProAlaProIylSerThrAlaProPro 634
    :::::
2059 GTATCGGACACACAGCTGCACAGATATGCTGCTGACACACAGATAT 2108
    :::::
635 AlaIsgLylValThrSerAlaPro.....AspThr..... 644
    :::::
2109 GTCGCTGACACACATGACGAGATATGCTGCGCGACACA.....C 2152
    :::::
645 .....AlaProAlaProIylSerThrAlaPro 654
    :::::
2153 ACACACAGTGTGCACAGATATGCTGTGACACGCGCGACACACATGACA 2202
    :::::
654 roAlaIsgLylValThrSer.....AlaProAspThrArg 665
    :::::
2203 TATGCTGCTGACACACATTCGACACAGCTGCGACAGGCGCGCAATAT 2252
    :::::
666 ProAla..ProIylSerThrAlaProPro.....AlaIsgLylValThs 680
    :::::
2253 GTCGCTGCGACACAGC.....GGATATGCTGTAGTACACACA 2293
    :::::
680 exAlaProAspThrArgProAlaProIyl..SerThrAlaProProAlaI 696
    :::::
2294 CACGCGACATGCTGTCCGACACACAC..... 2321
    :::::
696 sglValThrSerAlaProAspThrArgProAlaProIylSerThrAla 713
    :::::
2322 .....ACGATGCGACAGATATGCTGTCCGCGACACACACGCGCGAGAT 2366
    :::::

```

```

713  rorProAlaHisGlyValThrSerAlaProAspThrArgPro..... 726
2367  ATGTCGCTGGACACACACAGATATGCT.....GCCCAACACTCA 2410
727  ...AlaProGlySerThr..AlaProProAlaHisGlyValThrSerAlaP 742
2411  CACACGTGACAGATATGCTGGACACACACTGTCACAGATATGCTGTC 2460
742  rorAspThrArgProAlaProGlySerThrAlaProProAlaHisGlyVal 758
2461  TGCACATGACACACACAGTGCAGATATGCTGCCGATACACAGCAGCAC 2510
759  ThrSer.....AlaProAspThrArgPro..... 766
2511  ACATGACGATATGCTGCTGGACACACTTCCGACACACATGACACACA 2560
767  .....AlaProGlySerThrAlaProProAlaHisGlyValT 779
2561  CAGGTGCA...GATATGCTGCTGGACACACAGCAGACTGCTGTTTG 2607
779  hrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAla 795
2608  GGAGGTGCTGCCGTGAAGCCCTGACAGTGTGCTGAGGCTGATAGTT 2657
796  HisGlyValThrSerAlaProAspThrArgProAlaProGlySer.... 810
2658  GATGAGGACTTTCCTGCTGCACCGTCACT.....CCCCAAGCTGTC 2701
811  ....ThrAlaProProAlaHisGlyValThrSerAlaProAspThrArgP 826
2702  CCGCCTGTGTCGCCGCGCAGTCCCGCC.....TCATCACC 2739
826  rorAlaProGlySerThrAlaProProAlaHisGlyValThrSerAlaPro 842
2740  GCGCTGTGCTGCCCTGGCCGTATTTTTCACCTGCTGCTGGTGC 2789
843  AspThrArgProAlaProGlySerThr...AlaProProAlaHisGlyVa 858
2790  CCAGGACTCCCTTACTGCTGTGGCTGGGCTGGGCGACACAGCCCCA 2839
858  lThrSerAlaPro.....AspThrArgProA 867
2840  AGCCTGAGAGGCTGAGCCCATGCTAGTGGCTATGCCCACTCATCT 2889
867  lArProGlySerThrAlaProProAlaHisGly.....ValThrSer 880
2890  CCCCCGACACAGAGAGGCGCTGATTTATTTAAGAAATGAAGA 2939
881  AlaProAspThrArgProAlaPro..... 888
2940  TAAATATTAAATGAATGAGGAAGACTG..... 2969
889  .....GlySerThrAlaProProAlaHisGlyValThrS 900
2970  .GTTGACGAGACTGTGCTCTCTCGGCGCCGAGCCGCT..... 3011
900  erAlaProAspThrArgProAlaProGlySerThrAlaProAlaHis 916
3012  GGTCTTCAGGATGCTGATGACACACCCGCTGAGGCGGACACACACC 3061
917  GlyValThrSerAlaProAspThrArgProAlaProGlySerThrAlaP 933
3062  CCCC 3065
933  opfo 934
seq_name: SwissProt_39:NC_A2_HUMAN
seq_documentation_block:
AC NC_A2_HUMAN STANDARD; PRT; 761 AA.
PI3592; PI3593;
DT 01-JAN-1990 (Rel. 13, Created)
FT 01-APR-1990 (Rel. 14, Last sequence update)

```

```

DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEURAL CELL ADHESION MOLECULE, PHOSPHATIDYLINOSITOL-LINKED ISOFORM
DE PRECURSOR (N-CAM 120) (NCAM-120) (CD56 ANTIGEN).
GN NCAM1 OR NCAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBL_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (N-CAM 120).
RC TISSUE-Skeletal muscle;
RX MEDLINE=89305258; PubMed=3253057;
RA Barton C.H., Dickson G., Gower H.J., Rowett L.H., Putt W.,
RA Elsom V., Moore S.E., Goridis C., Walsh F.S.;
RT "Complete sequence and in vitro expression of a tissue-specific
RL phosphatidylinositol-linked N-CAM isoform from skeletal muscle."
RL Development 104:165-173(1988).
RN [2]
RP SEQUENCE OF 491-761 FROM N.A. (N-CAM 120).
RC TISSUE-Skeletal muscle;
RX MEDLINE=87301755; PubMed=2887295;
RA Dickson G., Gower H.J., Barton C.H., Prentice H.M., Elsom V.L.,
RA Moore S.E., Cox R.D., Quinn C., Putt W., Walsh F.S.;
RT "Human muscle neural cell adhesion molecule (N-CAM): identification
RL of a muscle-specific sequence in the extracellular domain."
RL Cell 50:1119-1130(1987).
RN [3]
RP SEQUENCE OF 491-655 FROM N.A. (SECRETED ISOFORM).
RX MEDLINE=89077552; PubMed=3203385;
RA Gower H.J., Barton C.H., Elsom V.L., Thompson J., Moore S.E.,
RA Dickson G., Walsh F.S.;
RT "Alternative splicing generates a secreted form of N-CAM in muscle
RL and brain."
RL Cell 55:955-964(1988).
CC -1- FUNCTION: THIS PROTEIN IS A CELL ADHESION MOLECULE INVOLVED IN
CC NEURON-NEURON ADHESION, NEURITE FASCICULATION, OUTGROWTH OF
CC NEURITES, ETC
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
CC -1- ALTERNATIVE PRODUCTS: THE DIFFERENT TISSUE-SPECIFIC FORMS OF
CC N-CAM ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -1- DATABASE: NAME-PROT; NOTE-CD guide CD56 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prot/cd/cd56.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X16841; CAA34739.1; -
DR EMBL: M17409; AAS59912.1; -
DR EMBL: M22094; AAS59910.1; -
DR EMBL: M22092; AAS59911.1; -
DR EMBL: M22091; AAS59911.1; JOINED.
DR PIR: S07784; ITHUNG.
DR PIR: A31635; A31635.
DR MIM: I16930; -
DR InterPro: IPR001777; -
DR InterPro: IPR003006; -
DR Pfam: PF00041; fn3; 2.
DR Pfam: PF00047; lg; 5.
KW Immunoglobulin domain; Cell adhesion; Glycoprotein; Repeat; Signal;
KW GPI-anchor; Alternative splicing.
FT SIGNAL 1 19
FT CHAIN 20 761 NEURAL CELL ADHESION MOLECULE,
FT DOMAIN 34 103 PHOSPHATIDYLINOSITOL-LINKED ISOFORM.
FT DOMAIN 132 196 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 228 294 IG-LIKE C2-TYPE DOMAIN.

```

```

FT DOMAIN 322 392 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 419 486 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 518 595 FIBRONECTIN TYPE-III.
FT DOMAIN 660 727 FIBRONECTIN TYPE-III.
FT DISULFID 41 96 PROBABLE.
FT DISULFID 139 189 PROBABLE.
FT DISULFID 235 287 PROBABLE.
FT DISULFID 329 385 PROBABLE.
FT DISULFID 426 479 PROBABLE.
FT CARBOHYD 222 222 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 315 315 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 449 449 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 478 478 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 635 655 GEPSAPKLEGMGEDNSIKV -> NIAONHCMMFOAGLH
NALMK (IN SECRETED ISOFORM).
SQ SEQUENCE 761 AA; 83770 MM; FOCAD3292D7AB67E CRC64;

```

```

alignment_scores:
  Quality: 284.50      Length: 559
  Ratio: 1.086         Gaps: 21
  Percent Similarity: 46.869   Percent Identity: 22.719

```

```

alignment_block:
US-09-598-042a-2 x NCA2_HUMAN ..

```

```

Align seg 1/1 to: NCA2_HUMAN from: 1 to: 761

```

```

103 GCCCGCTGGCGCCGACATGCGCTGCAAGTCCCAAGTGGAGGGAGCC 152
    |||:|||||:|||||: ||| ||| :||| ||| ||
225 Alaslnleuylglinservalthrlleuvalcysapalaglulphr 241
153 GCCCGCTGACCATGTGACCAAGATGCCACCACATCAAGCGGCT 202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
241 OGUpIrothmetsertrphrllysaprglulglulglug 258
203 GAGGCCCTTCGGGCTGCTG.....CCGAGGGCTGAAGTG 240
258 luspaspgrlulysytrilrpheserapsrserserglulethrlle 274
241 AAGCAGGTGAGCGGAGAGATGCCGGCTGACGTGCAAGGCCACCAA 290
    |||:|||||:|||||: ||| ||| :||| ||| ||
275 lylsyalaglyglulglulglulglulglulglulglulglul 291
291 CGGCTTGGCAGCTGAGCTCAACTACACCTCTGCTGCTGGATGACA 340
    |||:|||||:|||||: ||| ||| :||| ||| ||
291 nlysalaglyglulglulglulglulglulglulglulglul 304
341 TTAGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 390
304 ..... 304
391 GAGAGCCCCCGCCAGCAGAGTGGGACAGCCGCTTCACACAGCCCTC 440
    :|||:|||||:|||||: ||| ||| :||| ||| ||
305 .....Phealalysprolysilethrltyrvalgl 314
441 CAAGATGAGCGCGGGTATTCGACAGCGCGGTGAGTCCGCTGGCGC 490
314 uasnrgln.....Thralametgluleglulglulglulglul 327
491 TCAAGTGTGTCGCGAGCGGACCCCTGGCCGACATCAGCTGGATGAG 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
327 eulnrcysglualaserclylasprrlroserllethttrpargthr 343
541 GACGACCAAGGCTTGACGCCGCCAGAG.....GC 569
    :|||:|||||:|||||: ||| ||| :||| ||| ||
344 serthrlragsanlleserserglulglulglulglulglulglul 360
570 CGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 619
    :|||:|||||:|||||: ||| ||| :||| ||| ||
360 lvaltrgerhisalatrvalserleuthrlleuylserlletgint 377

```

```

620 CGGAGACAGGGGCAATACACCTGCCGCTGCAAGCCGGCGGGCC 669
    |||:|||||:|||||: ||| ||| :||| ||| ||
377 ytrthaspalaglyglulglulglulglulglulglulglulglul 393
670 ATCAAGCCACCTACAGAGTGGATGTATCCAGGACCGCTCCAGCC 719
    :|||:|||||:|||||: ||| ||| :||| ||| ||
394 aspsrgerlnsermettyrleuval.....gintyrvalabr 406
720 CGTGTCTACAGGACGACCCCGTGAACACGAGGTGAGCTGGGGGGA 769
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
406 Olyslleuengly.....Provalalavaltyrtrpgrlulglul 421
770 CCAGCTCTCCAGTCAAGGTGCGCAGCAGAGTGAACCGGTGATCCAG 819
421 lvalalasllethrcysglulglulglulglulglulglulglul 437
820 TGGCTGAAGCGCGGTGAGTACGCGCCGAGGCGCCACAACTCCACAT 869
    |||:|||||:|||||: ||| ||| :||| ||| ||
438 trppearq..... 440
870 CGATGTGGCGGCGCAGAGTTGTGTGTGCTGCCAGGGGTGAC..... 912
    |||:|||||:|||||: ||| ||| :||| ||| ||
441 .....Aspglyln.....Leuileuproserserasnlyrsera 452
913 .....GTGTGTGCGGCGCGCAGCGGCTCTACCTCAATAGCTGCTC 954
    :|||:|||||:|||||: ||| ||| :||| ||| ||
452 snlleyslletyrasntrpseralaserlyrleu..... 465
955 ATACCCCGTGGCGCCGAGCAGATGCGGCAATGATCTGCTGGGCGC 1004
    :|||:|||||:|||||: ||| ||| :||| ||| ||
466 valtrpaspserclulasnaspheglyasnlyrascysthralava 482
1005 CAACACCATGAGCTACAGCTTCGCA..... 1030
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
482 lasnaglllelglulglulglulglulglulglulglulglul 498
1031 GCGCTTCTCACTGCTGCGCAGACCAACCGCCAGGGCCACTGTG 1080
    :|||:|||||:|||||: ||| ||| :||| ||| ||
499 thrproserproserleasprlnvalglu..... 509
1081 GCCTCTCTCTCTCGGCGCAGTACGCTGCGCGGTGCGGTGATGAGCAT 1130
    |||:|||||:|||||: ||| ||| :||| ||| ||
510 .....Protyrserserthra 515
1131 CCCAGCGCGGCTGTCTTCATCTGCGGACCTGTCTGTGTGCTTGGC 1180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
515 lagln.....Valglnpheasr 520
1181 AGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1230
    :|||:|||||:|||||: ||| ||| :||| ||| ||
521 gluproglulalthrglulglulglulglulglulglulglulglul 533
1231 CACCGCGCGCGGAGAGCGCGGAGCGGAGCGGAGAGAGAGAGAG 1280
    :|||:|||||:|||||: ||| ||| :||| ||| ||
534 .....Lysalaglutpralglalavalglulglulglulglul 547
1281 CTGTTGGCGGCTTCAGCGCTGCGCTGTGCGCTGTGAGAGAG 1330
    :|||:|||||:|||||: ||| ||| :||| ||| ||
547 erlystrprylaspralalysglulalasermetlulglulglul 563
1331 ATGGGTCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1377
    :|||:|||||:|||||: ||| ||| :||| ||| ||
564 valgllyleuyls...Proglutntrthrltyralavalatrgleuallalale 579
1378 GCTGGCCTTAAGTTGTACCCAACTCTACACAGACATCCACACACAG 1427
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
579 uasnrglylsglyleuylglulglulglulglulglulglulglul 596
1428 ACACTGATTGCGCGCGCTGTGTGAGAGACATGAGTCTCGGACGCC 1477
    :||| ||| ||| ||| ||| ||| ||| ||| ||| |||
596 lnprovalhisserproproPro...ProalaseralaserSerthr 611

```

```

1478 CCACCACTTACTGGGCGCCGACGTTGCTGACCTTAAGTTGACCCCA 1527
||||| :||| :|||
612 ProValProLeuSerProIAspThrThrPro.....LeuProAl 626
1528 AACCTTACACAGATCCACACACACACACACTCTCACACACAC 1577
||||| :|||
626 aLeuAlaThrThrIuPro..... 632
1578 TCACACGTCGAGGCAAGTCCACACACATCCATCATGCTGAC 1627
||||| :||| :|||
633 .....AlaLysGlyIuProSerAlaPro.....LysLeuGlu 643
1628 GCACACGCTATCTGCAGAGCGCGGGCGCCGACAGACAGCAGACTG 1677
||||| :||| :|||
644 GlyIuMetGlyGluAspGlyAsnSerIleLysValAsnLeuIleLysG 660
1678 GCAGAGTGCAGAGCGAGCTGCAGA 1702
||||| :|||
660 nAspAspGlyGlySerProIleArg 668

seq_name: SwissProt_39:YHLL_EBV

seq_documentation_block:
ID YHLL_EBV STANDARD: PRT: 660 AA.
AC P03181:
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOHETICAL BHLFI PROTEIN.
OC Epstein-Barr virus (strain 995-8) (Human herpesvirus 4).
OC Viruses: dsDNA viruses, no RNA stage: Herpesviridae;
OC Gammaherpesvirinae: Lymphocryptovirus.
OX NCBI_TaxID=10377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84270667; PubMed=6087149;
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA Tufnell P.S., Barrett B.G.;
RT DNA sequence and expression of the 995-8 Epstein-Barr virus genome.;
RC Nature 310:207-211(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL, V01555; -; NOT_ANNOTATED_CDS.
DR PIR: A03742; Q08E3.
KW Hypothetical protein; Early protein; Repeat.
FT DOMAIN 149 648 4 x 125 AA TANDEM REPEATS.
FT REPEAT 149 273 1.
FT REPEAT 274 398 2.
FT REPEAT 399 523 3.
FT REPEAT 524 648 4.
SQ SEQUENCE 660 AA; 66244 MW; 86DA1D67A37152A2 CRC64;

alignment_scores:
Quality: 283.00 Length: 694
Ratio: 0.956 Gaps: 50
Percent Similarity: 42.651 Percent Identity: 29.827

alignment_block:
US-09-598-042A-2/rev x YHLL_EBV ..
Align seg 1/1 to: YHLL_EBV from: 1 to: 660
1773 GGGGCTATGCTGCACACACAGACTGCTGGGCTGCAGCTTCCTCCTC 1724
||||| |||
||||| |||

```

```

65 GLyAlaArgProGlyGlyLysAsnArgValGlyAlaGlyArg...GlyAr 80
1723 GCGATGGTCCCTCCCTGCTGCTGACAGCTCCGCTCCATCCTCCAGT 1674
||||| :||| :|||
80 gProGlyThrProAla.....ProSerArgGlns 90
1673 CTGCCCTGCTGGCGCCCGCCGTCGCTGACATACGATACGCTGCTG 1624
||||| :||| :|||
90 eArgArgThrGly.....ProAlaGluGlnAlaAspHisAlaHisSer 104
1623 AGCACTGATAGTGATG..... 1606
105 AsnProThrGlyGlyCysSerAspProGlnArgSerProArgThrArgG 121
||||| :||| :|||
1605 .....GCTGGT.....GCAC 1596
121 nAlaGlyThrAlaLeuGlyGlyGlySerAlaGlyLeuGlySerArgGlyP 138
||||| :||| :|||
1595 CTGGCCCT...CCAGCTGTGAGTGTGTGAGAGTGTGTGTGTGTGTG 1549
||||| :||| :|||
138 rArgProHisProAlaPheGlnValGlnTrpSerAlaArgAsnProGly 154
||||| :||| :|||
1548 TGTGATGCTGTGTGATGATGTTGGGTACAACTTAGGCCCAACTGTGG 1499
||||| :||| :|||
155 CysPro.....ArgThrTriArgArgArgSerGlyAlaGlnArgG 168
1498 CCTGGGCCCACTAGTCTGTGGGGGCTGCCGAGACCCATGCTCTCACA 1449
||||| :||| :|||
168 yHisProProGlyAlaGlyGlnArgPro.....SerGlyProt 182
1448 CAGGCGGCGCAATACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1399
||||| :||| :|||
182 hGlyGlyArgProAlaAlaProGlyAlaPro..... 192
1398 GGGGTACAACTTAGGCCCACTAGTGGCTGGGCCCACTAGTGTGTGG 1349
||||| :||| :|||
193 GlyThrProAlaAlaProGlyProGlyGlyAlaAlaAlaProSerG 209
1348 GGGTGGCCGAGAGCCATGCTCTCACAACGCCCAACGAGGCCACG 1299
||||| :||| :|||
209 yAlaThrProHisProGluArgGlySerGlyProAlaAspProAla 225
1298 CTGAGGGGG.....CCAGCGAGGAGTCTGTGTC..... 1265
||||| :||| :|||
226 AlaAlaArgLeuProProGluArgGlnGlnProArgLeuProGlnAsp 242
1264 .....CGCTGCGTGGCGGGCGCTCCCGCGG...GGCGGT 1232
||||| :||| :|||
242 euAlaAlaAlaGlnArgCysProAlaGlyProProThrArgSerGly 258
1231 GCCCAGCGAGGGGAGGGCGAGCGCGG.....GGTGCACGGCTTCTTC 1188
||||| :||| :|||
259 AlaAlaAlaGlnArgThrHisArgArgProProGlyCys..... 271
1187 TGGGCTGTGGCAAGCCACAGGAGCGAGGTCGCCAGATGAGAGACGCC 1138
||||| :||| :|||
272 ....ProArgSerAlaArgAsnProGlyCysProArgThrTrpArgAda 287
1137 GCGTGGGATGC...CGATGACCAGCGGCGCGCAGCTAGTGGCGAGG 1091
||||| :||| :|||
287 rGserGlyAlaGlnArgGlyHisProProProGlyAlaGlyGlnArgPr 303
1090 ACGAGGAGGCCACAGGTGCG...CCTGGCGGTTTGGTGTGGCAGCAG 1044
||||| :||| :|||
303 oSerGlyProThrGlyGlyArgProAlaAlaProGlyAlaProGlyThr 319
1043 GTGAGGAAGGGCTGGCGAAGCTGATGACCATGATGTTGGGCCCAAGCA 994
||||| :||| :|||
319 ..... 319
993 GATGTACATGCCCGCATGCTCTGCGGCGACGGGTGATGACAGCTTAT 944
||||| :||| :|||
320 .....ProAlaAlaProGlyProGlyGly..... 327

```



CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; M28231; AAA28728.1; ALT\_SEQ.

DR EMBL; X/6243; CAAS3822.1; -.

DR PIR; A32579; A32579.

DR PDB; 1CFB; 30-NOV-94.

DR FLYBASE; FBgn0002968; Nrg.

DR InterPro; IPR001777; -.

DR InterPro; IPR003006; -.

DR Pfam; PF00047; fn3; 5.

DR Pfam; PF00047; 1g; 6.

KW Cell adhesion; Glycoprotein; Transmembrane; Repeat; 3D-structure; Immunoglobulin domain; Signal; Embryo; Alternative splicing.

KW SIGNAL

FT CHAIN 1 23

FT DOMAIN 24 1239

FT TRANSMEM 24 1138

FT DOMAIN 1139 1154

FT DOMAIN 1155 1239

FT DOMAIN 53 123

FT DOMAIN 149 224

FT DOMAIN 262 329

FT DOMAIN 354 422

FT DOMAIN 447 512

FT DOMAIN 536 606

FT DOMAIN 629 690

FT DOMAIN 729 792

FT DOMAIN 832 896

FT DOMAIN 932 997

FT DOMAIN 1024 1098

FT DISULFID 59 111

FT CARBOHYD 625 706

FT CARBOHYD 182 182

FT CARBOHYD 198 198

FT CARBOHYD 411 411

FT CARBOHYD 448 448

FT CARBOHYD 652 652

FT CARBOHYD 683 683

FT CARBOHYD 821 821

FT CARBOHYD 1125 1125

FT CONFLICT 1234 1234

FT CONFLICT 1237 1237

SO SEQUENCE 1239 AA; 138284 MW; 49E12692DD0D27D CRC64;

Alignment\_scores:

Quality: 274.00 Length: 365

Ratio: 1.497 Gaps: 14

Percent Similarity: 50.137 Percent Identity: 25.205

Alignment\_block:

US-09-598-042a-2 x NRG\_DROME ..

Align seg 1/1 to: NRG\_DROME from: 1 to: 1239

76 CCGGTCGACGATTT.....CGTGGGACGTGGCCCGGCTGGCCG 116

|||||..... |||||..... |||||

246 ProValArgIntyValserArgArgInsereLeuAlaLeuArgIly 262

117 CASATGGGGTCGAGTGGCCGAGTGGAGGGAGCCGCCGCCGTCGACCA 166

262 SArgmetGluLeuPheCysIleTyrGlyLeuThrProLeuProGlnThr 279

167 TGTGAGCAGGATGAGCGGACACATGACAGGCGGTGAGCGGCTTCCGC 216

279 aItrPserIlyAspIlyGlnArgIleGln.....TrrPser.....Asp 291

217 GTGTCGCGGACGGG.....CTGAAGGTGAACAGAGTGA 251

292 ArgIleThrGlnGlnIlyHisTyrGlyLeuLysSerLeuValIleArgGlnThr 308

252 GCGGAGAGTGGCGGCTGATAGTGTGCAAGGCCACCAAGGCTTCGCA 301

..... |||||

308 nPheaspAspAlaGlyThrTyrThrCysAspValSerAsnGlyValGly 325

302 GC.....CTGAGGTCACACTACACCTCGTCGG 330

325 snAlaGlnSerPheSerIleIleLeuAsnValAsn..... 336

331 CTGATGATGACATTAGCCAGGAGGAGGAGGACCTGGGGCCGACAGCTCTC 380

336 ..... 336

381 TGGGGGTCAAGAGACCCGCCAGCAGCATGGGACAGCCGCGCTTCA 430

337 .....SerValProTyrPheT 342

431 CACAGCCCTCCAGATGAGGCGCGGGGTGATGACAGGCGCCGTGGGTAC 480

342 hTyrGlnThrProL.....IleAlaThrAlaAlaGlnAspGln 354

481 TCCGTCGCGGCTCAAGTGGCGTGGCCAGCGGACCCCTGGGCGGACATGC 530

355 GluValValPheGlyCysArgAlaAlaGlyValProGlnProLysIleSe 371

531 GTGATGAGAGGACGACAGGCTTGACGCGCCAGAGGCGCGCTGAGCCA 580

371 rTrpIleHisAsnGlyLysProIle.....GluGlnSerThrPro 385

581 GGAGAGAGAGTGG.....ACATGAGCCTGAAGACCTGGCG 618

385 snProArgArgThrValThrAspAsnThrIleArgIleHisLeuVal 401

619 CCGGAGAGACAGGCGCAATACACCTGCGCGGTGAGAACCGCGCGCGCC 668

402 LysGlyAspThrGlyAsnTyrGlyCysAsnAlaThrAsnSerLeuGly 418

669 CATCAAGCGCACCTCAAGGTGATGTATCCAGGAGACCGCTTCAAGC 718

418 rVal.....TyrLys...AspValTyrLeuAsnValGlnAlaGlnP 431

719 CCGTGTCTCAGGACGACCGCCGTGAACACGAGCGGTGACTGGGGGG 768

431 rOProThrIleSerGlnAlaProAlaAlaValSerThrValAspGlyArg 447

769 ACCAGTCTCTTCCAGTGCAGAGTGGCGGACGACGTGAACCCGCTATCA 818

448 AsnValThrIleLysCysArgValAsnGlySerProLysProLeuVal 464

819 GTGGCTGAAGCGCGTGAATACGAGCGCGGCGGCCGACATCCACCA 868

464 sTrpLeuArgAlaSerAsnTrpLeuThrGlyLeuArgTyrAsn..... 478

869 TCGATGTGGGCGGCCAGAAATTGTGTGTCTGCTGCCACGAGGTGACGTGG 918

479 .....ValGlnAlaAsnGlyAsp..... 484

919 TCGCGGCGGACGCGCTCTACCTCAATAAGCTGCTCATGACCCGTCGCG 968

485 .....LeuGlnIleGlnAspValThr 491

969 CCAGGACGATGCGGCGATGATACATGCTGCTGGCGGACACACATGAGGT 1018

491 rPheSerAspAlaGlyLysTyrThrCysTyrAlaGlnAsnLysPheGlyG 508

1019 ACAGCTTCGCGAGCGCTTCTACCGCTGCTGCCAGGACCCAA..... 1062

508 LuIleGlnAlaAspArgLysSerLeuValValLysGlnHisThrArgLeuThr 524

1063 .....CGGCAAGGCGACCTGTGGCTCTGCTGCTGCGGCACT 1101

525 GlnGlnProGlnAsnTyrGlyValAlaAlaGlyGlnSerAlaThr 539

seq\_name: SwissProt\_39:FGRL\_DROME

seq\_documentation\_block:







FT CONFLICT 992 992 S -> F (IN REF. 2)  
SQ SEQUENCE 1070 AA; 118260 MW; 47CDF25B8E3698A5 CRC64;

## alignment\_scores:

Quality: 272.00 Length: 530  
Ratio: 1.097 Gaps: 23  
Percent Similarity: 46.792 Percent Identity: 23.208

## alignment\_block:

US-09-598-042a-2 x PRK7\_HUMAN ..

Align seg 1/1 to: PRK7\_HUMAN from: 1 to: 1070

```
43 GCGGAGATTGCTTCATCCACGATGATGCTCCCGGCTGCAGATTCG 92
|||||
386 AlaglyalYThrCysHisAlaAlaSnLeuAlaglyLn.....Ar 400
|||||
93 TCGGCAAGTGGCCGGCTG..... 111
|||||
400 gATgGlnAspAlaSnlleThrValAlaThrValProSerTrpLeuLysL 417
|||||
112 .....GCGCGACATATGCGCTGCAGTGC 135
|||||
417 ysProGlnAspSerGlnLeuGlnGlyLysProGlyTyrrLeuAspCys 433
|||||
136 CCAAGTGGAGGGAGCCCGCGCTGACCATGTGACCAAGATGGCCG 185
|||||
434 LeuThrlGlnAlaThrProLysProThrValValTrpTyrrArgAsnGlnMe 450
|||||
186 CACCATCCACAGCGGCTGAGACCGCTCCGCTGCTCCGACAGGG...C 232
|||||
450 tLeuIle...SerGlnAspSerArgPheGlnValAlaPheLysAsnGlyThrl 466
|||||
233 TGAAGTGAACAGCATGTGAGCGGAGGATCCGCGCTGACGTGCAAG 282
|||||
466 euArgIleAsnSerValGluValTyrrAspGlyThrTrpTyrrArgCysMet 482
|||||
283 GCCACCAACGCTTCGGCAGCGCTGACGATGACAGACCGCTGCGTGC 332
|||||
483 SerSerThrProAlaGlySerIleGlnAlaGlnAlaValLeuGlnValLe 499
|||||
333 GGATGACATTAGCCCAAGGAGAGAGAGCTGGGCGCCGACAGCTCCTCTG 382
|||||
499 uGluLysLeu..... 502
|||||
383 GGGGTCAAGAGAGACCCCGCAGCAGCATGTGGCAGACCGCGTTACA 432
|||||
503 .....LysPheThr 505
|||||
433 CAGCCTCCAGATGAGCGCGGGGTGATGACAGCGCCCTGGGTAGCTC 482
|||||
506 ProProProGlnProGlnGln.....CysMetGlyPheAspLysG 519
|||||
483 CGTGGCGCTCAAGTGCCTGGCCAGCGGACCGCTGGCGGACATCACT 532
|||||
519 uAlaThrValProCysSerAlaThrGlyArgGlnLysProThrlLeuLys 536
|||||
533 GGAGGAGGAGACCGACGCGCTTCAGCGCGCCAGAGCGCGGTGACCCAG 582
|||||
536 TrpGluArgAlaAspLysSer...SerLeuProGlu..... 546
|||||
583 AAGAAGAGTGG.....ACACTGAGCCTGAAGAACT 614
|||||
547 .....TrpValThrAspAsnAlaGlyThrlLeuHisPheAlaArgVa 560
|||||
615 GCGGCGGAGAGACGCGCAATACACTGCGCGGTGTCGAA...CGCG 661
|||||
560 lThrArgAspAspAlaGlyAsnTyrrCysIleAlaSerAsnLysProG 577
|||||
662 CGGGCGCATCAAGCGCCACCTACAAAGTGGATGTGATCAAGCGGACCGGT 711
|||||
577 lngLysGlnlleArgAlaHis.....ValGlnLeuThrVal 588
```

```
712 TCACAGCCCGTGTCTACAGGACGACCCCGTGAACAGCAGCGTGGACTT 761
|||||
589 AlaValAlaPheIleThrPheLysValGluProGluArgThrThrValTyrG 605
|||||
762 CCGGGGACACACAGCTTCCTCACTGCAAGTGGCGGACGACGTGAAGCGG 811
|||||
605 nGlnHisThrlAlaLeuLeuGlnCysGluAlaGlnGlnLysProLysProL 622
|||||
812 TCATCCAGTGGCTGAAGCGCGTGGAGTACGGCGCGGACGCGCCCAAC 861
|||||
622 euIleGlnTrp..... 625
|||||
862 TCACACATGATGTGGCGGCGCAGAGATTGTGTGTGCTGCCACCGAGTGA 911
|||||
626 .....LysGlyLysAspArgIleLeuAspProThr...Ly 936
|||||
912 CGTGTGTGCGCGCGCGAGCGGCTCTACCTCAATAGCTGTATACCC 961
|||||
636 sLeuGlyProArgMetHisIlePheGlnAsnGlySerLeuValIleHisA 953
|||||
962 GTGCCCGCCAGACGATGGCGGCGATGTACATGCGCTGGCGGCGCAACCC 1011
|||||
653 sPValAlaProGlnAspSerGlyArgTyrrCysIleAlaLysAsnSer 669
|||||
1012 ATGGGCTACAGCTCCGACGCGCTTCCTCAACGCTGTGCTGCAGACCAAA 1061
|||||
670 CysAsnIleLysHisThrlGlnAlaProLeuTyrrValAlaAspLysProVa 686
|||||
1062 ACCG.....CGAGGCGCACCTGTGGCTCTCTGCTCTCGG 1096
|||||
686 lProGlnGlnSerGlnGlylProGlySerPro..... 697
|||||
1097 CCACATGCTGCGTGGCGCGCTGTGTC.....ATCGCATCCACCGCGC 1140
|||||
698 .....ProTyrrLysMetIleGlnThrlleGlyLeuSerValGly 710
|||||
1141 GCT.....GTCTCATCTGTGGGACACCTG...CTCTGTGCTTTGCCA 1181
|||||
711 AlaAlaValAlaTyrrIleAlaValAlaValLeuGlyLeuMetPheTyrrCys.. 726
|||||
1182 GCGCCAGAAAGACCGCTGCACCCCGCGCTGCGCTCCCGTCCGCGGCGG 1231
|||||
727 .....LysLysArgCysLysAlaLysArgLeuGlnLysGlnPro... 739
|||||
1232 ACCGCGCGCGGGGAGCGCGCGCGACCGGACGAGACAGACACTTCCC 1281
|||||
740 .....GluGlyGlnLysPro 744
|||||
1282 TCGTGGCGCGCTTCAGCGCTGCGCTGGTGGGCTGTGAGAGACA 1331
|||||
745 GluMetGlnCysLeuAsnGlyGlyPro.....LeuGlns 756
|||||
1332 TGGGTCCGCGCAGCCCGCAGCAC.....TTCAGTG 1363
|||||
756 nGlyGlnProSerAlaGlnIleGlnGlnGlnValAlaLeuThrlSerLeuG 773
|||||
1364 GCCAGGCGCAGTGTGCGCTTAAGTGTACCCCAACTATACAGAC 1413
|||||
773 LysArgGlyProAlaAlaThrlAsnLysArgHisSer.....Thrser 786
|||||
1414 ATCCACACACACACACTATATTGGGCGCGCTGTGTGAGAGAGCATGGG 1463
|||||
787 AspLysMetHisPheProArgSerSerLeuGlnProIleThrlThrlLeuG 803
|||||
1464 TCTCCGCGACCGCCCGCACTTACTGAGCGCGCGCCAG 1503
|||||
803 LysSerGlnPheGlyGlnValPheLeuAlaLysAlaGln 816
```

seq\_name: SwissProt\_39:KMLS\_HUMAN

seq\_documentation\_block:

ID KMLS\_HUMAN STANDARD; PRT: 1914 AA.

AC 015746; Q14844; Q16794; 095796; 095797; 095798; 095799; 09UB65;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DE MYOSIN LIGHT CHAIN KINASE, SMOOTH MUSCLE AND NON-MUSCLE ISOZYMES  
 DE (EC 2.7.1.117) (MLCK) [CONTAINS: TELOKIN (KINASE RELATED PROTEIN)  
 DE (KRP)].  
 GN MYLK OR MLCK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=umbilical vein endothelial cells;  
 RX MEDLINE=97304466; PubMed=9160829;  
 RA Garcia J.G.N., Lazar V., Gilbert-McClain L.I., Gallagher P.J.,  
 RA Verin A.D.;  
 RT "Myosin light chain kinase in endothelium: molecular cloning and  
 RT regulation";  
 RL Am. J. Respir. Cell Mol. Biol. 16:489-494(1997).  
 RN [2]  
 RP REVISIONS.  
 RA Birukov K.G., Garcia J.G.N.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORMS 2; 3A; 3B AND 4).  
 RC TISSUE=umbilical vein;  
 RX MEDLINE=99216419; PubMed=10198165;  
 RA Lazar V.L., Garcia J.G.N.;  
 RT "A single human myosin light chain kinase gene (MLCK; MYLK).";  
 RL Genomics 57:256-267(1999).  
 RN [4]  
 RP REVISIONS TO ISOFORM 2.  
 RA Birukov K.G., Garcia J.G.N.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 923-1914 FROM N.A.  
 RC TISSUE=hippocampus;  
 RX MEDLINE=96121365; PubMed=8575746;  
 RA Potler M.-C., Chelot E., Pekarsky Y., Gardiner K., Rossler J.,  
 RA Turnell W.G.;  
 RT "The human myosin light chain kinase (MLCK) from hippocampus:  
 RT cloning, sequencing, expression, and localization to 3qcen-q21.";  
 RL Genomics 29:562-570(1995).  
 RN [6]  
 RP SEQUENCE OF 1614-1914 FROM N.A.  
 RC TISSUE=Lung and Placenta;  
 RX MEDLINE=20007838; PubMed=10536370;  
 RA Watterson D.M., Schavocky J.P., Guo L., Weiss C., Chlenski A.,  
 RA Shtrik V.P., Van Eldik L.J., Halech J.;  
 RT "Analysis of the kinase-related protein gene found at human chromosome  
 RT 3q21 in a multi-gene cluster: organization, expression, alternative  
 RT splicing and polymorphic marker.";  
 RL J. Cell. Biochem. 75:481-491(1999).  
 RN [7]  
 RP SEQUENCE OF 1456-1914 FROM N.A.  
 RC TISSUE=Placenta;  
 RA Watterson M.D.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: CALCIUM/CALMODULIN-DEPENDENT ENZYME RESPONSIBLE FOR  
 CC SMOOTH MUSCLE CONTRACTION VIA PHOSPHORYLATION OF A SPECIFIC SERINE  
 CC IN THE N-TERMINUS OF MYOSIN LIGHT CHAINS (MLC), AN EVENT THAT  
 CC FACILITATES MYOSIN INTERACTION WITH ACTIN FILAMENTS. CENTRAL  
 CC DETERMINANT IN THE DEVELOPMENT OF VASCULAR PERMEABILITY AND TISSUE  
 CC EDEMA FORMATION. IN THE NERVOUS SYSTEM IT HAS BEEN SHOWN TO  
 CC CONTROL THE GROWTH INITIATION OF ASTROCYTIC PROCESSES IN CULTURE  
 CC AND TO PARTICIPATE IN TRANSMITTER RELEASE AT SYNAPSES FORMED  
 CC BETWEEN CULTURED SYMPATHETIC GANGLION CELLS. CRITICAL PARTICIPANT  
 CC IN SIGNALING SEQUENCES THAT RESULT IN FIBROBLAST APOPTOSIS.  
 CC -1- CATALYTIC ACTIVITY: ATP + [MYOSIN LIGHT-CHAIN] -> ADP + [MYOSIN  
 CC LIGHT-CHAIN] PHOSPHATE.  
 CC -1- SUBUNIT: TELOKIN BINDS CALMODULIN.

CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS OF MLCK ARE PRODUCED  
 CC BY ALTERNATIVE INITIATION. A NON-MUSCLE FORM (WHICH IS THE LONGEST  
 CC FORM); A SMOOTH-MUSCLE FORM AND TELOKIN (A C-TERMINAL SECTION WITH  
 CC NO CATALYTIC ACTIVITY).  
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 6 ISOFORMS: 1 (SHOWN HERE), 2, 3A,  
 CC 3B, 4 AND DEL-1790; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: SMOOTH MUSCLE AND NON-MUSCLE ISOZYMES ARE  
 CC EXPRESSED IN A WIDE VARIETY OF ADULT AND FETAL TISSUES AND IN  
 CC CULTURED ENDOTHELIUM WITH QUALITATIVE EXPRESSION APPEARING TO BE  
 CC NEITHER TISSUE- NOR DEVELOPMENT-SPECIFIC. NON-MUSCLE ISOFORM 2 IS  
 CC THE DOMINANT SPLICED VARIANT EXPRESSED IN VARIOUS TISSUES. TELOKIN  
 CC HAS BEEN FOUND IN A WIDE VARIETY OF ADULT AND FETAL TISSUES.  
 CC -1- PTM: MLCK IS PROBABLY DOWN REGULATED BY PHOSPHORYLATION.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL: U48959; AAC18423.2; -;  
 CC DR EMBL: AF069601; AAD15921.2; -;  
 CC DR EMBL: AF069602; AAD15922.1; -;  
 CC DR EMBL: AF069603; AAD15923.1; -;  
 CC DR EMBL: AF069604; AAD15924.1; -;  
 CC DR EMBL: X85337; CA559685.1; -;  
 CC DR EMBL: AF096771; AAD51380.1; -;  
 CC DR EMBL: AF096766; AAD51380.1; JOINED.  
 CC DR EMBL: AF096767; AAD51380.1; JOINED.  
 CC DR EMBL: AF096768; AAD51380.1; JOINED.  
 CC DR EMBL: AF096769; AAD51380.1; JOINED.  
 CC DR EMBL: AF096770; AAD51380.1; JOINED.  
 CC DR EMBL: AF096774; AAD54018.1; -;  
 CC DR EMBL: AF096771; AAD51381.1; -;  
 CC DR EMBL: AF096769; AAD51381.1; JOINED.  
 CC DR EMBL: AF096770; AAD51381.1; JOINED.  
 CC DR EMBL: X90870; CAA62378.1; -;  
 CC DR HSP: P56276; ITLK.  
 CC DR MIM: 600922; -;  
 CC DR InterPro: IPR000719; -;  
 CC DR InterPro: IPR001777; -;  
 CC DR InterPro: IPR002290; -;  
 CC DR InterPro: IPR003006; -;  
 CC DR Pfam: PF00041; Tn3; 1.  
 CC DR Pfam: PF00047; Ig; 8.  
 CC DR Pfam: PF00069; Pkinase; 1.  
 CC DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 CC DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 CC KW Transferase: Serine/threonine-protein kinase; Calmodulin-binding;  
 CC KW ATP-binding; Phosphorylation; Immunoglobulin domain; Repeat;  
 CC KW Alternative initiation; Alternative splicing.  
 CC FT CHAIN 1 1914  
 CC FT ISOYME.  
 CC FT CHAIN 923 1914  
 CC FT MYOSIN LIGHT CHAIN KINASE, SMOOTH-MUSCLE  
 CC FT ISOYME.  
 CC FT CHAIN 1761 1914  
 CC FT TELOKIN.  
 CC FT INIT\_MET 923 923  
 CC FT FOR MYOSIN LIGHT CHAIN KINASE, SMOOTH-  
 CC FT MUSCLE ISOZYME.  
 CC FT INIT\_MET 1761 1761  
 CC FT FOR TELOKIN.  
 CC FT DOMAIN 1343 1413  
 CC FT FIBRONECTIN TYPE-III.  
 CC FT DOMAIN 1464 1719  
 CC FT PROTEIN KINASE.  
 CC FT DOMAIN 1711 1774  
 CC FT CALMODULIN-BINDING.  
 CC FT DOMAIN 1824 1891  
 CC FT IG-LIKE C2-TYPE DOMAIN.  
 CC FT NF\_BIND 1470 1478  
 CC FT ATP (BY SIMILARITY).  
 CC FT BINDING 1493 1493  
 CC FT ATP (BY SIMILARITY).  
 CC FT ACT\_SITE 1585 1585  
 CC FT BY SIMILARITY.  
 CC FT DOMAIN 1906 1914  
 CC FT POLY-GLU.

```

FT DOMAIN 868 998 5 X 28 AA APPROXIMATE TANDEM REPEATS.
FT REPEAT 868 995 I-1.
FT REPEAT 896 923 I-2.
FT REPEAT 924 951 I-3.
FT REPEAT 952 979 I-4.
FT REPEAT 980 998 I-5 (INCOMPLETE).
FT DOMAIN 999 1063 6 X 12 AA APPROXIMATE TANDEM REPEATS.
FT REPEAT 999 1063 II-1 (INCOMPLETE).
FT REPEAT 1004 1015 II-2.
FT REPEAT 1016 1027 II-3.
FT REPEAT 1028 1039 II-4.
FT REPEAT 1040 1051 II-5.
FT REPEAT 1052 1063 II-6.
FT VARSPLC 437 506 VSGIPKPEVAMELESTPVPROEGSIEVEYDAGSHYCLILKA
FT REPEAT 1433 1439 DEVEUSD -> MKMRCOT (IN ISOFORM 3A,
FT VARSPLC 1473 1546 ISOFORM 3B AND ISOFORM 4).
FT REPEAT 1473 1546 GKFQGVFLVETKTKRKVAGKFKFVSAKREKENTROETISIM
FT VARSPLC 1555 1705 NCLHHPKLVOCVDAFEERKANIVMLEIVSGEL -> L
FT REPEAT 1790 1790 (IN ISOFORM 4).
FT CONFLICT 933 933 MISSING (IN ISOFORM 3A AND ISOFORM 3B).
FT CONFLICT 963 963 V -> M (IN REF. 5).
FT CONFLICT 1022 1022 S -> P (IN AAD15922).
FT CONFLICT 1048 1050 P -> A (IN REF. 5).
FT CONFLICT 1162 1162 KPM -> EAH (IN REF. 5).
FT CONFLICT 1210 1210 P -> L (IN AAD15922 AND AAD15923).
FT CONFLICT 1280 1280 L -> P (IN REF. 5).
FT CONFLICT 1284 1284 E -> D (IN AAD15922 AND AAD15923).
FT CONFLICT 1300 1300 M -> I (IN AAD15922, AAD15923 AND
FT CONFLICT 1316 1316 AAD15924).
FT CONFLICT 1326 1326 A -> G (IN REF. 5).
FT CONFLICT 1478 1478 L -> S (IN REF. 5).
FT CONFLICT 1511 1511 T -> S (IN REF. 5).
FT CONFLICT 1563 1563 V -> C (IN REF. 5).
FT CONFLICT 1609 1609 S -> T (IN AAD15922 AND AAD15923).
FT CONFLICT 1639 1639 I -> T (IN REF. 5).
FT CONFLICT 1640 1640 A -> P (IN REF. 5).
FT CONFLICT 1648 1648 G -> R (IN REF. 5).
FT CONFLICT 1659 1659 LS -> PF (IN REF. 5).
FT CONFLICT 1710 1711 CT -> LA (IN REF. 5).
FT CONFLICT 1897 1897 L -> H (IN AAD15922, AAD15923 AND
SO SEQUENCE 1914 AA; 210773 MW; 9295FE0279CB0033 CRC64;

```

```

alignment_scores:
  Quality: 267.50      Length: 825
  Ratio: 0.727        Gaps: 42
  Percent Similarity: 44.606  Percent Identity: 21.818

```

alignment\_block:

US-09-598-042a-2 x KMLS\_HUMAN ..

Align seg 1/1 to: KMLS\_HUMAN from: 1 to: 1914

```

112 GGGCGACATATGCGGCTGCAGTCCAGTGAGGGGACCCGCCGCCCT 161
113 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
47 GYLAthlthAlaYsPhegluglYArgValaRglYtYrProgluProgl 63
162 GAGCAGTGTGACCAAGATGGCGCACATCCAGCAGGGCTGTGAGCGGCT 211
63 nAlthrTPrlhIsArGsmelYglInProletlrnSerGly...GlyArGp 79
212 TCCGCGTGTGTCGCGCAGGGCTGAAG.....GTGAAG 243
79 he...LeuLeuAspGysglYlleArgglYThrPheSerleuValIleHis 94
244 CAGGTGAGCGGGAGGATGCCGGCGTGTACGTGTGCAGGCCACCAACG 293

```

```

95 AlAlValhIscluclunspargglYsTYrThrCysgluAlaThrAsnGl 111
294 CTTCGGCAGCCTGAGCTGACACTACACCCCTGCTGCTGTGATGACATTA 343
111 YserglYAlaArGlnValThrValgluThrValgluGlYserPheA 128
344 GCCCAGAGG.....AAGGAGAGCCTGGGCCCGCAGCAGC 375
128 lAlYsGlInleuGlInProValValSerlYsThrleuGlY...AspArg 143
376 TCCTGTGGGGGTCAAGAGACCCCGCAGCAGCAGCTGGGACGA..... 420
144 PheSerAlaSerAlaValaJgluThrArgProSerlYleTropolYglucyspr 160
421 .CGGCCCTTCACACACCCCTCCCAAGATGAGCGCCCGGTATGACACGCG 469
160 opPolYsPheAla.....ThrlYsleuGlYArgValaValYsGlU. 174
470 CGGTGGGTAGCTCGTGGCTCAAGTGCAGTGGCGCAGCGGACCTCGG 519
175 ....GlylnmetglYArgPheSerCyslYsThrThrglYArgProglIn 189
520 CCGCAGATCAGTGTGATGAAGAGCAGCAGCAGCCTTGACGCGCCAGAGGC 569
190 ProglInValThrTrpLeuYsGlyAsnValProleu...GlnProSerAl 205
570 CGGTGAGCCCGAGAGAGAGAG.....TGGACACTGAGCCTGAAGA 610
205 ArArgValSerValSerGlulYsAsnGlYmetGlInValleuGlulIleHis 222
611 ACCTGGCGCGGAGAGCAGCAGCGCAATACACTCGCGCTGTGCAACCGC 660
222 lYAlAsnGlInAspAspValglYValTYrThrCysleuValaValAsnGlY 238
661 GCGGGCGCATCAAGCCACCTCAAGAGTGTGATGATTCACGCGGACCGG 710
239 SerGlYlYsAlaSerMetSerAlaGlulSerlYle..... 250
711 TTCGAGCCCGGTCTCAAGAGCAGCAGCAGCCCGTGAACAGCAGGTGAGCT 760
251 .....GlnlYleuAspSerAlaAsnArGserPhe...V 261
761 TCGGGGAGACACAGCTCTTCAGTGCAGGTGGCGCAGCAGCAGTGAACCG 810
261 AlArggluThrlYsAlaThrAsnSerArpValaRglYsGlUValThrAsn 277
811 GTGATC.....CAGTGGCTGAAGCGGTGGAGTACGGCCCGCAGG 851
278 ValIleSerlYsGlulSerlYsleuAspSerlYleuGlulAlaAlaLysSe 294
852 CCGCACACATCCACCATGATGATGGGGCGCGCAGAGTTGTGTGCTGCTG 901
294 rLysAsnCYsSerSerProglInArgglYlYser.....P 306
902 CCAGGGGTGACGTGTGTCGCGCGCGCAGCGGCTCTACCTCAATAAGCTG 951
306 robroTrrpAlaAlaAsnSerGlInProglInProArArglUserlYsleu 322
952 CTGATCAACCGTCCCGCCAGCAGCAGTGGCGCATGTACATCTGCGCTGG 1001
323 GluSerCYsLysAspSerProArGThrAlaProglInThrProValleuGl 339
1002 CGGCACACATGAGGCTTACAGCTTCGCGAGCGGCTTCACACGCTGCTG 1051
339 nLysThrSerSerSerlYleThrleuGlInAlaAla.....ArgValGlInp 354
1052 CAGACCCAAACCCGCGAGGCGCACCTGTGGCTCTGTCTGTCGCGCACAT 1101
354 rogluProArGAlaProglYleuGlYValleuSer..... 365
1102 AGCCTCGCGTGGCGGTGATGAGCATGCCATCCAGCGGCGCTGTCTTCAAT 1151

```

```

366 .....Prosergly..... 368
1152 CCTGGGACACCTGCTCTGCTGGCTTTGGACGGCCGACAGAGAACCGCTGCA 1201
369 .....GluGluArgLysArgProAlaP 376
1202 CCCCCGCGCTCCCTCCCTGCTGGCGACCGCCGCGCGGAGCGGC 1251
376 roProArgProAla...ThrPhePro...ThrArgGlnProGlyLeu... 389
1253 CGGACCGCGACGGGAGACAGACCTTCCCTGCTGGCGCGCTGACGCC 1301
390 .....GlySerGlnAspValSerLysAlaAlaAsnArgArg 402
1302 TGGCCCTGCTGGGG..... 1317
402 gllleProMetGluGlnArgAspSerAlaPheProLysPheGluSerL 419
1318 .....CTGGGT 1323
419 ysProGlnSerGlnGlnValLysGlnAsnGlnThrValLysPheArgCys 435
1324 GAGGAGACATGGGTCTCCGCGACGCCCGAC.....CACTTACTGGGCC 1367
436 GluValSerGlyIlePro...LysProGluValAlaIlePheLeuGlu 451
1368 AGCCCACTGCTGGCCCTAACTTGTACCCCAACTCTTACACAGACATCC 1417
451 yThrProValArgArgGlnGluSerIleGluValTyrgLysAlaArg 468
1418 ACACACACACACACTGTATTGGGGCGCTGTGGAGGAGCATGGGTGC 1467
468 LySerHisIleuLysLeuLysAlaArgThrArgAspSerGly... 483
1468 CGGACGCCCGCCCACTTACTGGGCCCGACGGCCAGTTGCTGGCCCTAAG 1517
484 .....ThyTyr.....Se 486
1518 TTGTACCCCAACTCTTACACAGATCCACACACACACACACACTC 1567
486 rCysThrAlaSerAsnAlaGlnGlyGlnVal...SerCysSerThrPheLeu 502
1568 TCACACACACTCACACGTGGAGGCAAGSTCCACACAGACATCCACTATC 1617
503 Gln.....ValGluArgLeuAl 508
1618 AGTGTAGACGGCACCGTATCTGCAGAGGGGACGGGGGGCGCCGACAG 1667
508 aValMetGluValAlaProSerPheSerSerValLeuLysAspCysAlaV 525
1668 AGGCAGACTGGAGGATGGAGGACGAGCTGCAGACGAAGGACGGGAGCC 1717
525 alIleGluGlnGlnAspPheValLeuGlnCysSerValArg...GlyThrPr 541
1718 CATGGCGAGG.....AGAAATGGCAGACCCCGACGAG... 1751
541 ovalProArgIleThrTrpLeuAsnIleGln...ProIleGlnTyra 557
1752 .....TCTGTGTGAGGACATAGCCCTGGACACACACACACACACAC 1796
557 laArgSerThrCysGlnAlaGlyValAlaGlnLeuHisIleGln...Asp 572
1797 ACACACTGCGATGATGTATGCACACACATGGCGGCACACAGCTGCTCCT 1846
573 AlaLeuProGluAspHisGlyThrTyThrCys..... 583
1847 GAAGGCACACGTAAGCACACACAGCATGCACAGATATGCCCTGGGCA 1896
583 ..... 583
1897 CACAGATAGCTGCCAATATGACGACGACACAGACAGACATGCCAGAAC 1946
584 .....LeuAlaGluAsn 587

```

```

1947 ATACAGACATGCTGCTGACACATACACAGCAGACACCATGGCGAGATG 1996
588 AlaLeuGlyGlnValSer.....Cys 594
1997 TCTGCTGGACACACACACACAGCATATGCTGTGACGACACACA 2046
594 sSerAlaIlePheValThrValHisGluLys.....LysSerSerArgL 608
2047 CGTCGAGATATGCTATC.....CGGACACACAGTCGCACAGAT 2084
608 ySerSerIleuTrpLeuProValAlaProSerLysProThrAlaProIle 624
2085 ATGCTGCTGGACACACA.....GATATGCTGCTGAC 2119
625 PheLeuGlnGlyLeuSerAspLeuLysValMetAspGlySerGlnValTh 641
2120 ACACACATGC.....ACGATATGCTGCTGACACAC 2148
641 rMetThrValGlnValSerGlyAsnProProProGluValIleThrLeuH 658
2149 ACACACACACACGCTGTGCACAGATATGCTGTGACACGACACACATG 2198
658 sAsnGlyLysGlnIleGlnGluSerGlu..... 667
2199 CAGATATGCTGCTGGACACACACTTCGACACACAGCTGACAGGCGCAG 2248
668 .....AspPheHisPheGluGln....ArgGlyThrI 677
2249 ATATGCTGCTGGACACACGCGATATGCTGTAGTCACACACACACAGC 2298
677 nHisSerLeuTrpIleGln...GluValPheProGluAspThrGlyThrT 693
2299 AGACATGC.....TGTCCGACACACACACAGCAGCAG 2333
693 yThrCysGlnAlaIleTrpAsnSerAlaGlyGluValArgThrGlnAlaVal 709
2334 ATATGCTGCTGGACACACAC 2354
710 LeuThrValGlnGluProHis 716

seq_name: SwissProt_39:AXOL_HUMAN
seq_documentation_block:
ID AXOL_HUMAN STANDARD; PRT; 1040 AA.
AC Q02246;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE AXONIN-1 PRECURSOR (AXONAL GLYCOPROTEIN TAG-1) (TRANSIENT AXONAL
DE GLYCOPROTEIN 1).
GN TAXI OR TAG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI-TaxID=9606;
RX MEDLINE=93145965; Pubmed=8425542;
RA Hasler T.H., Rader C., Stoeckli E.T., Zuelliig R.A., Sonderegger P.;
RT "CDNA cloning, structural features, and eucaryotic expression of
RT human TAG-1/axonin-1.";
RN Eur. J. Biochem. 211:329-339(1993).
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=94140354; Pubmed=8307567;
RA Tsiotra C.P., Karagozeos D., Theodorakis K., Michaelidis M.T.,
RA Modi W.S., Furlley J.A., Jessel M.T., Papamtheakis J.;
RT "Isolation of the cDNA and chromosomal localization of the gene
RT (TAG1) encoding the human axonal glycoprotein TAG-1.";
RL Genomics 18:562-567(1993).

```

```

CC -1- FUNCTION: MAY PLAY A ROLE IN THE INITIAL GROWTH AND GUIDANCE OF
CC AXONS. MAY BE INVOLVED IN CELL ADHESION.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE NEURONAL MEMBRANE BY A
CC GPI-ANCHOR AND IS ALSO RELEASED FROM NEURONS.
CC -1- SIMILARITY: CONTAINS 6 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 4 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X68274; CAA4835.1; -
DR EMBL: X67734; CAA47963.1; -
DR PIR: S28830; S28830.
DR MIM: 150197; -
DR InterPro: IPR001777; -
DR InterPro: IPR003006; -
DR Pfam: PF00041; fn3; <4.
DR Pfam: PF00047; 1g; 6.
KW Immunoglobulin domain; glycoprotein; signal; GPI-anchor;
KW Cell adhesion; Repeat.
FT SIGNAL 1 28
FT CHAIN 29 1012
FT PROPEP 1013 1040 AXONIN-1.
FT DOMAIN 54 118 REMOVED IN MATURE FORM (POTENTIAL).
FT DOMAIN 148 216 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 254 313 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 341 402 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 433 495 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 523 594 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 606 612 GLY/PRO-RICH.
FT DOMAIN 611 706 FIBRONECTIN TYPE-III.
FT DOMAIN 714 809 FIBRONECTIN TYPE-III.
FT DOMAIN 816 908 FIBRONECTIN TYPE-III.
FT DOMAIN 917 1003 FIBRONECTIN TYPE-III.
FT SITE 744 796 CELL ATTACHMENT SITE (BY SIMILARITY).
FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 204 244 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 477 477 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 488 498 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 525 525 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 830 830 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 918 918 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 940 940 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 1012 1012 GPI-ANCHOR (POTENTIAL).
SQ SEQUENCE 1040 AA; 113393 MW; 254E78D3C28EBF6 CRC64;

alignment_scores:
Quality: 266.00 Length: 611
Ratio: 0.996 Gaps: 22
Percent Similarity: 43.699 Percent Identity: 21.277

alignment_block:
us-09-598-042a-2 x AXOL_HUMAN ..
Align seg 1/1 to: AXOL_HUMAN from: 1 to: 1040

73 TTCCTGGGTCAGAGATTTCGTGGCAGGTGGCCGCGTGGCCGCACTAT 122
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
245 PheProAlaGluThrTyr.....AlaLeuValGlyGlnGlnVala 257
123 GCGGCTGCACATGCCCACTGAGGAGGAGACCGCGCGCTGACCATGGA 172
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
257 LThrLeuGlyCysPheAlaPheGlyAsnProValProArgIleLysTrpA 274
173 CCAAGGATGGCGCAGCATCCACAGCGGCTGAGACCGCTTCGGGTGCTG 222

```

```

274 rglYsValaLysGlySerLeuSerProGlnTrpThr.....ThraIa 287
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
223 CCGCAGGCGCTGAGGTGAAGTGAAGCAGGTGGACGGAGATCCGCGGTGA 272
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
288 GluProThrLeuGlnIleProSerValSerPheGlyAspIleLysThrTy 304
273 CGATGTCAGAGCCACCAAGGCTTCGGCAGCTGAGGCTGACATACACCC 322
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
304 rGlucYsGlnAlaGlnLysSerLeuLysGlyArgAspThrValGlnGlyArg 321
323 TCGTGTGCTGGATGATGATTTAGCCACAGGAAGAGACCTGGGCGCGAC 372
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
321 LeuLeuVal..... 323
373 AGCTCCTTGGGGGTCAAGAGACCCGCCACGACGATGGCGACGAC 422
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
324 .....GlnAlaGlnProGlnLysPhe..... 330
423 GCGGTTTCACACAGGCCCTCCAAAGATGAGGCGCGGGTGTGCA..... 465
331 .....LysValIleSerAspThrG 337
466 ..CGGCGCTGGGTAGCTCCGTCGCGCTCAAGTGCCTGAGCCAGCGGAC 513
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
337 LuAlaAspIleGlySerAsnLeuArgTrpGlyCysAlaIleAlaGlyLys 353
514 CCTGCGCCGCAATCAGCTGGATGAGAGACGACCGCTTGAAGCGGCC 563
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
354 ProArgProThrValArgTrpLeuArgAsnGlyGlyProLeuIleAsnArg 370
564 AGAGCGCGCTGAGCGCCAGGAAGAAGTGGACACTGAGCTGAGAAC 613
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
370 AsnArgValGln.....ValLeuAlaGlnAspLeuArgPheSerLysL 385
614 TGGCGCGGAGGAGACAGCGGCAAAATACACTGCGCGCTGAGAACCGGCG 663
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
385 euseLeuGlnAspSerGlyMetLysGlnCysValAlaGlnAsnLysHis 401
664 GCGGCGCATCAAGCCGCAAGTGAAGTGGATGATCAGCGGACCGCTTC 713
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
402 GlyThrIleLeuArgAlaSerAlaGlnLeuAlaVal.....GlnAl 414
714 CAAGCCGCTGCTCACAGGACGACGCCCTGTAACACAGCGGTGACTTC. 762
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
414 AlaLeuAlaProAspPheArgLeuAsnProValArgGlyLeuIleProAla 431
763 .....GGGGGACACAGCTCCTTCAGTGAAGTGGCGACGACGTTGAAG 807
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
431 IaArgGlyGlyGlnIleLeuIleProCysGlnProArgAlaIleProLys 447
808 CCGGTGATCCAGTGGCTGAAGCGCGGTGAGATACGCGCGAGGGCGCCA 857
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
448 AlaValValLeuThrPheSerLysGlyThrGlu..... 457
858 CAACTCCACCATGATGTGGCGCGCAGAGTTGTGTGCTGCCACGG 907
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
458 .....IleLeuValAsnSer 463
908 GTGACGTGTGTGGCGCGCGCAGCGCTCCTACCTCAATAAGCTGCTATC 957
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
463 eArgValThrValThrProAspGlyThr.....LeuIleIle 475
958 ACCCGTCCCGCCAGAGCAGATGCGGCGATTCATCTGCTTGCGCCAA 1007
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
476 ArgAsnIleSerArgSerAspGlnGlyLysTrpThrCysPheAlaGlnAs 492
1008 CACCATGGGCTACAGCTTCGCGACAGCGCTTCACCGCGCTGCGACACC 1057
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
492 nPheMetGlyLysAlaAsnSerThrGlyIleLeuSerValArgAspAla 509
1058 CAAACCGCGCAGGCGCACACTGTGTGCGCTCTGTCTTCGCGCACTAGCTG 1107
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```



```

FT DOMAIN 1275 1897 CYTOSOL PLASMIC (POTENTIAL).
FT DOMAIN 1360 1806 PROTEIN-TYROSINE PHOSPHATASE.
FT DOMAIN 1649 1897 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 1538 1538 BY SIMILARITY.
FT ACT_SITE 1829 1829 BY SIMILARITY.
FT CARBOHYD 107 107 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 711 711 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 956 956 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT MOTAGEN 1538 1538 C->S: LOSS OF ACTIVITY.
SQ SEQUENCE 1897 AA; 211844 MW; 439850FID5C031EF CMC64;

```

```

alignment_scores:
  Quality: 264.00 Length: 842
  Ratio: 0.756 Gaps: 41
Percent Similarity: 41.449 Percent Identity: 20.784

```

alignment\_block:

US-09-598-042A-2 x PTFP\_HUMAN

Align seg 1/1 to: PTFP\_HUMAN from: 1 to: 1897

```

22 ATGCTCCCGCGCGCGCGCGCGGAGTGC..... 51
4 LeuValProAlaLeuValMetLeuGluValAlaGlyAlaHisGlyAs 20
20 pserTyrProValPheIleLeuValPro.....G 30
52 .....TCCTTCATCCAGTCGATGTCGTCCTCCGGTCGACGATTC 91
92 GTCGCGCAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 141
30 LuaspGlnThrGlyLeuSerGlyGlyValAlaSerPheValCysGlnAla 46
142 GAGGGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 191
47 ThrIleGlnProLysProAlaThrIleThrIleThrIleThrIleThrIle 63
192 CCACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 225
63 LserSer.....GlnArgPheGluValIleGluPheAspGlyAlaG 78
226 .....CAGGGGCTGAAGTGAAGCAAGCGTGGAGCGGAG 258
78 LysSerValLeuArgIleGlnProLeuArgValGlnArg..... 90
259 GATCCCGCGCGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 308
91 AspGlnAlaIleIleThrIleGlyCysThrAlaThrAsnSerLeuGlyIleAs 107
309 CGTCACTACACCTCGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCG 352
107 nThrSerAlaLysLeuSerValLeuGlnGluGlnLeuProGlyLys 124
353 AGAGAGAC.....CTGGGGCGCGCAGACGCTCTGTGGGGGTCAAGAGAC 356
124 heProSerIleAspMetGlyPro..... 131
397 CCGCGCAGCCAGCAGTGGCGCAGCAGCGCTTCACACAGCCCTCAAGAT 446
131 ..... 131
447 GAGCGCGCGGTGATGCGACGCGCGCGGTGATGCTCGCGCGCGTCAAGT 496
132 GlnLeuLysValValGlnLys.....AlaArgThrAlaThrMetLeu 146
497 GCCTGGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 543
146 YsaIaIaIaGlyLysProAspProGlnIleSerIlePheLysAspPhe 162
544 .....GACCAAGCCTTGACGCGCCCAAGAGCGCGCGTGAAGCCAGAA 584

```

```

163 LeuProValAspProAlaThrSerAsnGlyArgIleLysGlnLeuArg.. 178
585 GAAGAAGTGGCACTGACGACCTCGAAGAACCTCGCGCGCGAGACAGCGGCA 634
179 .....SerGlyAlaLeuGlnIleGlnSerSerGlnGlnSerAspGlnIle 194
635 AATACACTGCGCGCGTGTGCGAACCGCGCGCGCGCGCGCGCGCGCGCG 684
194 YstYrGlyCysValAlaThrAsnSerIleGlyThr.....ArgYrSerAla 209
685 AAGGTGATGTGATTCACGAGCGGACCGCTTCACAGCCGCTGCTCACAGGAC 724
210 ProAlaAsnLeuThrValArgValAlaArgValAlaProAlaPheSerIle 226
735 GCACCCCGGTGACAGCAGCGGTGACCTTCGCGCGCGCGCGCGCGCGCG 784
226 eProPheSerSerGlnGlnValMetProGlyGlySerValAsnLeuThrC 243
785 GCAAGTGGCGCAGCAGCTGACGCGCGGTGATCCAGTGGCGTGAAGCGCGTG 834
243 YsaValAlaValAlaGlyAlaProMetProThrValAlaIleStrpMet..... 256
835 GAGTACGCGCGCGCGAGGCGCGCACAACTCCACCATGATGTCGCGCGCA 884
257 ...MetGlyAlaGlu..... 260
885 GAAGTTGTGTGTCGTCGCCCGCGGTGACGTCGTGTCGCGCGCGCGCGCT 934
261 .....GlnLeuThrLysGlnAspGln 268
935 CTTACTCATTAAGCTGCTCATCACCCTGCGCGCGCGCGCGCGCGCGCG 984
268 etProValGlyArgAsnValLeuGlnLeuSerAsnValAlaArgSerAla 284
985 ATGTACATCTGCTTGGCGCGCAACACCATGCGCTACAGCTTCGCGCGCG 1034
285 AsnTyrThrCysValAlaIleSerSerLeuGlyMetIleGlnAlaThrAl 301
1035 CTTCCTCAGCGTGTGCGCGCAAGCCCAACCGCGCGCGCGCGCGCGCGCT 1084
301 agnValThrValLysAlaLeuProLysProIleAspLeuValValT 318
1085 CTTGCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1113
318 hrcGlnThrThrAlaThrSerValThrLeuThrIlePheAspSerGlyAsnSer 334
1114 ...CCGCTGCTC...ATGCGCATCCAGCGCGCGCGCTGCTTCAATCCTGG 1157
335 GluProValThrTyrIleGlyLeuGlnTyrArgAlaIaIaGlyThrGlnG 351
1158 CACCC.....TGCTCCTGTGGC 1174
351 LysProheGlnGlnValAspGlyValAlaThrThrArgTyrSerIleGly 367
1175 TTTCGAGCGCGCAGAGAAGCGGTGACCGCGCGCGCGCGCGCGCGCG 1224
368 GlyLeuSerProPheSerGlnThrAlaPheArgValLeuAlaValAsnSe 384
1225 CMTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1274
384 rIleGly.....ArgGlyProProSerGlnAlaValAlaValAlaArgT 398
1275 CC.....TTCCCTGTTGGCGCGC..... 1294
398 hrcGlyGlnGlnAlaProSerSerProProArgValAlaIaIaArgMet 414
1295 .....TCAGCCTGCGCGCTGTGCGCGCTGTGAGAGAGCATGGG 1335
415 LeuSerAlaSerThrMetLeuValGlnTrp..... 424
1336 TCTCGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1385

```



```

425 .....GluProProGluGluProAsnGlyLeuValArgGlyTyrArgV 439
1386 TAAGTTGACCCCAAACTCTACACAGACATCCACACACACACACTGTA 1435
      |||||::|||
439 aLTYrTYrThrProAspSerArg.....GluTYrArgValTr 446
1436 TTGGGGCCGCTGTGTGAGAGACATGGGTCTCCGGCAGCCCCCAGCACT 1485
      |||||::|||
447 ..ArgProPro.....AsnAlaTrp..... 452
1486 TACTGGGCCAGGCCAGTGTGCTGGCCTAAGTTGTACCACCAACTTAC 1535
453 .....H 453
1536 ACAGACATCCACACACACACACACACTCTCAGAC..... 1573
      |||||::|||
453 lStyHis...AsnThrAspAlaGlyLeuLeuThrThrValGlySerLeu 468
1574 .....ACACTCACACGTGGAGC 1590
469 LeuProGlyIleThrTYrSerLeuArgValLeuAlaPheThrAlaValGly 485
1591 GCAAGGTCCACACACACATCCATTCAGTGTAGACGCGACCCGTATCTG 1640
      |||||::|||
485 yAspGlyProProSerProThrIleGlnVal.....L 496
1641 CAGAGGCGACGGGGGGCCGCCACACAGCGAGCTGGAGAGATGGAGCA 1690
      |||||::|||
496 ySThGlnGlnGlyAlProAlaGlnProAlaAspPheGlnAlaGlyVal 512
1691 CGGAGCTCAGACGAAAGCGAGGGGCCATGGCGAGGAGATGGCCAGC 1740
      |||||::|||
513 GluSerAspThrArgIleGln.....LeuSerTrpLeuLe 524
1741 ACCC..... 1744
524 uProProGlnGluArgIleIleMetTYrGluLeuValTYrTrpAlaIaG 541
1745 .....CAGGAGCTGTGTGTGAGGCAATGACCCCTGGACACACA 1783
      |||||::|||
541 LuAspGluAspGlnGlnHisLYsValThrPheAspProThrSerTyr 557
1784 CACACAGACACACACATACCTGATGCATGTATGCACACACATCGCGC 1833
      |||||::|||
557 rThrLeuGluAspLeuLYsProAspPThrLeuTYrArgPheGlnLeuAla 574
1834 ACACGTGCTCCCTGAAGCGACACACGTACGACACACATGCACAGATA 1883
      |||||::|||
574 LeArgSerAsp..... 577
1884 TGCCGCTGGGCGACACAGATAAGTCCCAATGCAGCGACACAGACA 1933
      |||||::|||
578 .....MetGlyValGlyValPheThrProThrIleGluAlaArgThr.. 591
1934 GACATGCCAGAAATACAAAGGACATGCTGCCGAACATACACACGACAC 1983
      |||||::|||
592 .....AlaGlnSerThrProSerAlaPro 600
1984 CCATGCGC...AGATGCTGCTGCTGACACACACACACAGGATATGC 2030
      |||||::|||
600 roGlnIlyValMetCysValSerMetGlySerThrThrValArgValSer 616
2031 TGCTTGAGCGCACACAG...TGCAGATATGGTATCCGACA..... 2069
      |||||::|||
617 TrpValProProProAlaAspSerArgAsnGlyValIleThrGlnTyrSe 633
2070 .....CACAGTGCACAGATATGCTGCTGGACAC..... 2099
      |||||::|||
633 rValAlaHisGluAlaValAsp.....GlyLYsAspArgGlyArgH 647
2100 .....ACAGATAATGCTGCTGACACACACATGCAGGATTTGSC... 2141
      |||||::|||
647 lSValAlaAspGlyLeuSerArgGlnHisSerSerTrpAspLeuValGly 663

```

```

2142 .....TGCACACACACACACACAGTGTGCACATATGCTGTCTG 2182
      |||||::|||
664 LeuGluLYsTrpThr.....GluTYrArgValTr 673
2183 GACAGCGACACACATGCAGATA 2204
      |||||::|||
673 pValArgAlaHisThrAspVal 680
seq_name: SwissProt_39:AXO1_RAT
seq_documentation_block:
ID AXO1_RAT STANDARD; PRT; 1040 AA.
AC P22063;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE AXONIN-1 PRECURSOR (AXONAL GLYCOPROTEIN TAG-1).
OS TAXI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 31-41.
RC TISSUE=Spinal cord;
RA Furely A.J., Morton S.B., Manalo D., Karagozeo D., Dodd J.,
RA Jessell T.M.;
RT "The axonal glycoprotein TAG-1 is an immunoglobulin superfamily
RT member with neurite outgrowth-promoting activity.";
RL Cell 61:157-170(1990).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE INITIAL GROWTH AND GUIDANCE OF
CC AXONS. MAY BE INVOLVED IN CELL ADHESION.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE NEURONAL MEMBRANE BY A
CC GPI-ANCHOR AND IS ALSO RELEASED FROM NEURONS.
CC -1- TISSUE SPECIFICITY: IN NEURAL TISSUES IN EMBRYOS, AND IN ADULT
CC BRAIN, SPINAL CORD AND CEREBELLUM.
CC -1- DEVELOPMENTAL STAGE: TRANSIENTLY EXPRESSED ON A SUBSET OF AXONS
CC IN THE DEVELOPING RAT NERVOUS SYSTEM.
CC -1- SIMILARITY: CONTAINS 6 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 4 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M31725; AAA42201.1; -
DR PIR: A34695; A34695.
DR InterPro: IPR001777; -
DR InterPro: IPR003006; -
DR Pfam: PF00047; fn3; 4.
DR Pfam: PF00047; Ig; 6.
KW Immunoglobulin domain; Glycoprotein; Signal; GPI-anchor;
KW Cell adhesion; Repeat.
FT SIGNAL 1 30
FT CHAIN 31 71015
FT PROPEP 71016 1040
FT DOMAIN 56 120
FT DOMAIN 150 218
FT DOMAIN 256 315
FT DOMAIN 343 404
FT DOMAIN 435 497
FT DOMAIN 525 596
FT DOMAIN 608 614
FT DOMAIN 613 708
FT DOMAIN 716 811
FT DOMAIN 818 910
FT DOMAIN 911 1005
AXONIN-1.
REMOVED IN MATURE FORM.
IG-LIKE C2-TYPE DOMAIN.
IG-LIKE C2-TYPE DOMAIN.
IG-LIKE C2-TYPE DOMAIN.
IG-LIKE C2-TYPE DOMAIN.
IG-LIKE C2-TYPE DOMAIN.
IG-LIKE C2-TYPE DOMAIN.
IG-LIKE C2-TYPE DOMAIN.
GLY/PRO-RICH.
FIBRONECTIN TYPE-III.
FIBRONECTIN TYPE-III.
FIBRONECTIN TYPE-III.
FIBRONECTIN TYPE-III.

```

FT SITE 796 798 CELL ATTACHMENT SITE (POTENTIAL).  
FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 479 479 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 500 500 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 527 527 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 777 777 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 832 832 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 920 920 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 942 942 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 1040 AA; 113042 MW; 6E707EF6614CB4FB CRC64;

alignment\_scores:  
Quality: 262.00 Length: 618  
Ratio: 0.996 Gaps: 23  
Percent Similarity: 42.557 Percent Identity: 22.006

alignment\_block:  
US-09-598-042a-2 x AXOL\_RAT

Align seg 1/1 to: AXOL\_RAT from: 1 to: 1040

73 TTCCGGGTCGACGATTTCTGCGCAGGTGCGCGCGCGCCTGAT 122  
||||| : : : : :  
247 PheProProGluThrTyrr.....AlaLeuValGlyGlnGlnIva 259  
123 GCGGCTGACGTGCGCAGTGCAGAGGGGAGCCCGCGCTGACCATGGA 172  
: : : : : : : : : : :  
259 LThrLeuGlnGlnPheGlnPheGlnPheGlnPheGlnPheGlnPhe 276  
173 CCAAGATGCGCGCAGCATCCACAGCGCGTGCAGCGCTTCCGCGTCTG 222  
||| : : : : :  
276 rGlyValAspGlySerLeuSerProGlnPheAla.....ThrAla 289  
223 CCGCAGGCGCTGAAGTGAAGCAGGTGAGCGGAGGATGCGCGCTGTA 272  
||||| : : : : :  
290 GluProThrLeuGlnIleProSerValSerPheGlnLysPheGlnTyrr 306  
273 CGTGTGAGCGCCACCAAGCGCTTCCGCGCAGCTGACCTACATCACCC 322  
| : : : : : : : : : : :  
306 rGlyGlyGlnIleGlnLysSerLysGlyArgAspThrValGlnIleYarg 323  
323 TCGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 372  
: : : : :  
323 LeuLeuVal..... 325  
373 AGCTCTCTGCGGGGTCAAGAGAGACCCCGCAGCAGCAGTGGCGACGAC 422  
||||| : : : : :  
326 .....GlnAlaGlnProGlnIlePheLeu..... 332  
423 GCGGTTCACACAGCCCTCCAAAGATGAGGCGCGCGGTATGCGCA..... 465  
333 .....LysValIleSerAspThrG 339  
466 .CGGCGCGGTAGCTCCGCGGCTCAAGTGCAGTGCAGCGCGGAC 513  
: : : : : : : : : : :  
339 LuAlaAspIleGlySerAsnLeuArgTrpGlySerAlaAlaIleGlyLys 355  
514 CTTGCGCGCCCATCATCAGTGGATGAGAGGAGCAGCGCTTGAAGCGGCC 563  
| : : : : : : : : : : :  
356 ProArgProMetValArgTrpLeuArgAsnGlyGlnProLeuAlaSerG 372  
564 AGAGGCGCGTGAAGCCCGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGA 613  
: : : : : : : : : : :  
372 nAsnArgValGln.....ValLeuAlaGlyAspLeuArgPheSerLys 387  
614 TCGCGCGCGGAGAGCAGCGCAATACACTGCGCGGTGTCGAGACCGCGG 663  
| : : : : : : : : : : :  
387 euserLeuGlnAspSerGlyMetTrpGlnGlyValAlaGlnAsnLysHis 403

664 GCGCGCATCAAGCCGACCTCAAGAGTGGATGTGATCCAGCGGACCGCTTC 713  
| : : : : : : : : : : :  
404 GlyThrIleLeuValAspAlaGlnLeuAlaVal.....GlnAl 416  
714 CAAAGCGCGTGTCAACAGGACGACCCCGTGAACACGACGCTGACTTC. 762  
: : : : : : : : : : :  
416 AlaLeuAlaProAspPheArgGlnAsnProValArgArgLeuIleProAla 433  
763 ....GGGGGACCACTCTTCATGTCGAAAGTGCAGCAGCAGCAGTGAAG 807  
| : : : : : : : : : : :  
433 IaArgGlyGlyGlnIleSerIleLeuGlySerGlnProAlaAlaIleProLys 449  
808 CCGGTATCCAGTGGCTGAAGCGCGGTGAGTACGCGCGCGGCGCGCA 857  
: : : : : : : : : : :  
450 AlaThrIleLeuThrSerLys.....GlyThrGlnIleLeuG 462  
858 CAATCCACCATCATGATGTGGCGCGCAGAAATTGTGTGCTGCCACG 907  
| : : : : :  
462 YasnSerThr..... 465  
908 GTGACGTGTGTGCGCGCGCGCGCGCTTCACTCAATGAAGTGTGATC 957  
| : : : : :  
466 .ArgValThrValThrSerAspGlyThr.....LeuIle 477  
958 ACCGTCGCGCGCAGGACGATCGCGGATGATGATGATGATGATGATGAT 1007  
: : : : : : : : : : :  
478 ArgAsnIleSerArgSerAspGlyGlyLysTrpThrCysPheAlaGln 494  
1008 CACCATGCGCTTACAGCTCCGACAGCGCTTCTACCGCTGTCGACGACC 1057  
| : : : : : : : : : : :  
494 nPheMetGlyLysAlaAsnSerThrGlyIleLeuSerValArgAspAla 511  
1058 CAAACCGCGCGGCGCACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1107  
||| : : : : :  
511 hIrys..... 512  
1108 CCGTGGCGCGGTGTCATCGCATCCAGCGCGCTGCTTCACTGCGG 1157  
| : : : : : : : : : : :  
513 .....IleThrLeuAlaProSerSerAlaAspLeuAlaG 525  
1158 CACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1207  
| : : : : : : : : : : :  
525 YAspAsnLeuThrLeuGlnCysHisAlaSerHisAspProThrMetAsp 542  
1208 CCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1239  
| : : : : : : : : : : :  
542 euthrPheThrTrpThrLeuAspAspPheProIleAspPheAspLysPro 558  
1240 CCGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1289  
| : : : : : : : : : : :  
559 GlyGlyHisTrpArgArgAlaSerAlaLysGlnThrIleGlyAspLeu 575  
1290 CCGCTGACGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1335  
: : : : : : : : : : :  
575 rIleLeuAsnAlaHis.....ValArgHisGlyL 586  
1336 .....TTCGCGCGCGCG 1347  
586 ystrThrCysMetAlaGlnThrValValAspGlyThrSerLysGlnAla 602  
1348 CCGCAGCACTTACTGGC...CCAGCGCGAGTGTGCTGCGCTAAGTTGTA 1394  
: : : : : : : : : : :  
603 ThrValLeuValArgIleProProGlyProProGlyGly..... 615  
1395 CCGCAACCTCTACAGACATCCACACACACA..... 1428  
: : : : : : : : : : :  
616 ...ValValValArgAspIleGlyAspThrThrValGlnLeuSerTrp 631  
1429 .....CACTGATGCGCGCGCGCGCTGTGAGAGACATG 1461  
631 eArgGlyPheAspAsnHisSerProIleAlaLys.....TyThrLeu 645  
1462 GGTCTCG 1511



```

397  LunetProLysAspSerGlnLeuGlnGluSerLysProGlyTyrLeuHis 413
352  .....AAGAGAGACCTGGGGCC..... 369
414  CysLeuSerLysAlaSerLeuLysProThrValThrTyrArgAsnG1 430
370  .....GACAGCTCCTCGGGGGT.... 387
430  ValSerLleSerGluAspSerArgPheGluLleSerGluAsnGlyThrL 447
387  ..... 387
447  euArgLleAsnAsnValGluValTyrAspGlyThrMetTyrLysCysVal 463
388  CAAGAGAGACCCCGCCAGC.....CAGCACTGGGCGAGAGCC..... 423
464  SerSerThrProAlaGlySerLleGluGlyTyrAlaArgValHisValLe 480
424  .....CGCTTCACACAGCCCTCCAGAGATGAGGCGCGGGGTATCG 463
480  uGluLysLeuLysPheThrProProGlnProLeuGln.....C 494
464  CAGGCGCCGTGGGTAGCTCCGTGGCTCAAGTGCCTGGCCAGCGGCAC 513
494  yMetGluPheAsnLysGluValThrValSerCysSerAlaThrGlyArg 510
514  CTTGGGCGCCCATCACTGATGAGAGAGACAGGCGCTGACGCGGCC 563
511  GluLysProThrLleGlnThrLysThrAspLysSerLeuProse 527
564  AGAGGCGCGTACGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 613
527  rHisValSerHis.....ArgAlaGlyLleLeuSerPheHisLysV 541
614  TGGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
541  AlSerArgSerAspSerGlyAsnLysThrLysLleAlaSerAsnSerPro 557
661  GCGGCGCCCATCAAGCCACCTACAGGTGATGATCCAGCGGACCG 710
558  GlnGlyLysLleArgAlaThrValGlnLeuValAlaValTyrValTh 574
711  TTCACAGCCCGTGTCTACAGAGCAGCAGCCCGGTGACACAGCGGTGACT 760
574  rPheLys.....LeuGluProGluProThrThrValTyrG 586
761  TCGGCGGAGACAGCTCCTCCAGTGAAGGTGGCGAGCGAGCGAGAGCG 810
586  LngLysHisThrAlaMetPheGlnCysGlnAlaGluGlyAspProValPro 602
811  GTGATCCAGTGTGAGAGCGGTGAGATACGCGCGCGCGCACAA 860
603  HisLleGlnThrP..... 606
861  CTCACCATCATGATGTGGCGCCAGAGTTGTG.....GTGC 898
607  .....LysGlyLysAspLysLleLeuAspProSerLysLeuL 619
899  TGGCCAGCGGTGAGCTGTGTGGCGCGGCTCCTACCTCAATAAAG 948
619  euProArg.....LleGlnLleMetProAsnGlySer..... 629
949  GTGCTCATCAACCGTGGCCGCGAGAGAGATGCGGCGATGATGCTTGCCT 998
630  LeuValLleTyrAspValThrThrGluAspSerGlyLysTyrThrCysI 646
999  TGGCGCCCAACACCTGGGCTACAGCTCCGAGCGCTTCTCAACGCTGC 1048
646  euAlaLysAsnSerCysAsnLleLysHisArgGluAlaPheLeuTyrValV 663
1049  TGGCAGAGCCA.....AAACCGCGCAGGCGCACCTGTGGCTCTCTCG 1089

```

```

663  aLAspLysProAlaAlaGlnGluAspGlnGlyPro..... 674
1090  TCGTGGCGCCATGACCTGGCCGCGCGGTGATCATGAGCATCCAGCGG 1139
675  SerSerHisThrProTyrLysMetLleGlnThrLleGlyLysSerValG1 691
1140  CGCT.....GTCTTCATCTGGGACACCTGCTCTGTGGC 1174
691  yAlaAlaValAlaTyrLleLleLleValLeuGly.....LeuMetPheT 706
1175  TTTGCCAGGCGCCAGAGAGAGCGGTGACACCCCGCGCTGCCCTCCCG 1224
706  yCysLysLysArgArgLys.....AlaLysArgLeu 716
1225  CTTGGGACCGCCCGCGGAGAGCGCGCGAGCCGAGCGAGAGACAGA 1274
717  LysLysHis.....ProGluGly.....G1 723
1275  CCTTCCCTGTTGGCGCCCTCAGCGCTGCG 1305
723  uGluProGluMetGluCysLeuAsnGlyGly 733

seq_name: SwissProt_39:NCAL_BOVIN
seq_documentation_block:
ID NCAL_BOVIN STANDARD; PRT; 853 AA.
AC P31836;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NEURAL CELL ADHESION MOLECULE, 140 KDA ISOFORM PRECURSOR (N-CAM 140)
DE (NCAM-140).
GN NCAM1 OR NCAM.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Brain cortex;
RX MEDLINE=89376239; PubMed=2776887;
RA Lipkin V.M., Khramtsov N.V., Andreeva S.G., Moshayakov M.V.,
RA Petukhova G.V., Rakitina J.V., Feshchenko E.A., Ishchenko K.A.,
RA Mirzoeva S.F., Chernova M.N., Drianytsyna S.M.;
RT "Calmodulin-independent bovine brain adenylylate cyclase. Amino acid
RT sequence and nucleotide sequence of the corresponding cDNA.";
RL FEBS Lett. 254:69-73(1989).
RN [2]
RP SEQUENCE OF 20-36.
RX MEDLINE=86140120; PubMed=3512556;
RA Rougon G., Marshak D.R.;
RT "Structural and immunological characterization of the amino-terminal
RT domain of mammalian neural cell adhesion molecules.";
RL J. Biol. Chem. 261:3396-3401(1986).
RN [3]
RP IDENTIFICATION AS N-CAM.
RX MEDLINE=92111748; PubMed=1765159;
RA Premont R.T.;
RT "A bovine brain cDNA purported to encode calmodulin-insensitive
RT adenylyl cyclase has extensive identity with neural cell adhesion
RT molecules (N-CAMs).";
RL FEBS Lett. 295:230-231(1991).
RN [4]
RT FUNCTION: THIS PROTEIN IS A CELL ADHESION MOLECULE INVOLVED IN
RN NEURON-NEURON ADHESION, NEURITE FASCICULATION, OUTGROWTH OF
RN NEURITES, ETC.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: THE DIFFERENT TISSUE-SPECIFIC FORMS OF
CC N-CAM ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE A CALMODULIN-
CC INDEPENDENT ADENYLYLATE CYCLASE.

```





```

FT CARBOHYD 4339 4339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4351 4351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4362 4362 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4373 4373 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4422 4422 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4438 4438 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4502 4502 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4616 4616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4627 4627 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4752 4752 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4787 4787 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4881 4881 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4888 4888 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4955 4955 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4970 4970 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 5019 5019 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 5038 5038 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 5069 5069 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 1351 1351 H -> L (IN REF. 3).
FT CONFLICT 1412 1412 T -> S (IN REF. 3).
FT CONFLICT 1449 1449 L -> P (IN REF. 3).
FT CONFLICT 1504 1504 M -> T (IN REF. 3).
FT CONFLICT 4192 4192 G -> S (IN REF. 2).
SQ SEQUENCE 5179 AA; 540295 MM; 85CD7571FB9A5663 CRC64;

```

```

alignment_scores:
  Quality: 260.50      Length: 1055
  Ratio: 0.588         Gaps: 52
  Percent Similarity: 41.991   Percent Identity: 23.318

```

alignment\_block:

US-09-598-042A-2 x MUC2\_HUMAN ..

Align seg 1/1 to: MUC2\_HUMAN from: 1 to: 5179

```

263 CCGGGGTAGCTGTGCAAGGCCACCAAGCGCTGGGACGCTGAGCGTC 312
    ||||| ||| ||||| |||||
3521 ProThrProThrGlyThrGlnThrProThr..... 3530
313 AACATACACCGCTGCTGCTGATGATATACCCAGGAGAGAGAGCCT 362
    ||||| ||| ||||| |||||
3531 ThrThrProThrThrThrThrThrValThr..... 3541
363 GGGGCGGCACA...GCTCTCTGGGGGTCAGAGAGACCCCGCCAGCAGC 409
    ||||| ||| ||||| |||||
3542 ....ProThrProThrProThrGlyThr...GlnThrProThrThrThr 3555
410 AGTGGGACAGACCGGCTTACACAGACCCCTCAAGATGAGGCGCGGTG 459
    ||||| ||| ||||| |||||
3556 ProThrThrThrThrThrValThrProThrProThrProThrGly... 3571
460 ATCCGACGCGCGGTAGTCCTCGCGGCTCAAGTGGCGGCGGAGCGG 509
    ||||| ||| ||||| |||||
3572 ThrGlnThrProThrThrThrThrProThrThrThrThrThrValThr 3588
510 GCACCTCTGGCGGCATCAGCTGATGAAGAGACAGACCGCTTGCAGCG 559
    ||||| ||| ||||| |||||
3588 roThrProThrProThrGlyThrGlnThrProThrThrThrThrThrThr 3604
560 GCCCAGAGGCGCGCTAGCCAGGAGAGAGAGAGTGCAGCTGAGCGTGAAG 609
    ||||| ||| ||||| |||||
3605 ThrThrThrThrThrValThrPro.....ThrProThr 3614
610 AACCTGCGCGCGGAGACAGCGCAATATACCTGCGCGGCTGTCAGAC. 658
    ||||| ||| ||||| |||||
3614 rProThrGlyThrGlnThrProThrThrThrThrThrThrThrThrThr 3631
659 .....GCGGCGCGCGCATCAACGCCCA 679
3631 hrValThrProThrProThrProThrGlyThrGlnThrProThrThrThr 3647
680 CCTACAGGTGATGTGATCCAGCGGACCGCTTCCAGCCCGTGTCTCACCA 729

```

```

3648 ProThrThrThrThrThrValThrProThrProThrProThr...ThrG 3663
    ||||| ||| ||||| |||||
730 GGCAGCAGACCCCGTACAGACAGAGGTGACTTCGGGGGAGCCAGCTGCTT 779
    ||||| ||| ||||| |||||
3663 yThrGlnThrProThrThr.....ThrPro 3672
780 CCAGTCAAGGTGCGCAGCGAGCGATGAGCGGTGATCCAGTGGCTGAGAAC 829
    ||||| ||| ||||| |||||
3672 LeThrThrThrThrThrValThrProThrProThrProThrGlyThrGln 3688
830 GCGGTGAGAGCGCGCGCGCGGCGCCCAACTCCACCATGATGGGCG 879
    ||||| ||| ||||| |||||
3689 ThrProThrThrThrThrThrThrThrThr..... 3699
880 GGCAGAGATTGTGTGTGCTGCCAGCGGGTACGTGTGCTCGCGGCCGA 929
    ||||| ||| ||||| |||||
3700 ....ThrValThrProThrProThrProThrGlyThrGlnThrProT 3714
930 CG.....GCTCTACCTCAATAGC 949
    ||||| ||| ||||| |||||
3714 hrThrThrProThrThrThrThrThrValThrProThrProThrProThr 3730
950 TCGCTATACCGCTGCGCGCGCGAGAGATGGGGATGTACATGTGCGCT 999
    ||||| ||| ||||| |||||
3731 ThrGlyThrGlnThrProThrThrThrThrProThrThrThrThrThrVa 3747
1000 GCGCGCAACACCATGAGGCGTACAGCTCCGACGCGCTGCTCCTC..... 1041
    ||||| ||| ||||| |||||
3747 LThrProThrProThrProThrGlyThrGlnThrProThrThrThrThrPro 3764
1042 .....ACGTGTGSCAGACCCCAAAACCGCAGGG..... 1071
    ||||| ||| ||||| |||||
3764 LeThrThrThrThrThrValThrProThrProThrProThrGlyThrGln 3780
1072 .....CCACGTGGGCTCGCTGCTGCGGCGACATGAGCTGCGC 1109
    ||||| ||| ||||| |||||
3781 ThrProThrThrThrThrProThrThrThrThrThrValThrProThrPr 3797
1110 GTGGCCCGTGTGTATCGGATCCAGCGCGCGCTGCTCATCTGTGGCA 1159
    ||||| ||| ||||| |||||
3797 oThrProThrGlyThrGlnThrProThrThrThrThrProThrThrThrThr 3814
1160 CCGTCTCTGTGCTTGTGCGAGCGCCAGAGAGACCGTGCACCCCGCG 1209
    ||||| ||| ||||| |||||
3814 hrThrVal.....ThrProThr 3819
1210 CCGTGGCCCTCGCTGCTGCGAGACGCGCGCGGGAGCGCGCGGAGCG 1259
    ||||| ||| ||||| |||||
3820 ProThrPro.....ThrGlyThrGlnThrProThrThrThrThrThrThr 3834
1260 CAGCGGAGACAAGACCTCCCTGCTGGCGCGCTCAGACGCTGCGCGCTG 1309
    ||||| ||| ||||| |||||
3834 hrThrThrThrThrThrValThrProThrProThrProThr..... 3846
1310 GTGTGGGCTGTGTGAGAGACATGGGTCTCGGACAGCCCGCCAGACATTA 1359
    ||||| ||| ||||| |||||
3847 .....GlyThrGlnThrProThrThrThrThrThrThrThrThr 3855
1360 CTGGGCGCCAGCGCCAGTTCCTGCGCCCTAGTGTACCCCAACTGTACAC 1409
    ||||| ||| ||||| |||||
3855 oLeThrThrThrThrThrValThrProThrProThrProThrGlyThrG 3872
1410 AGACATTCACA.....CACACACACATGTATGGCGCGCT 1447
    ||||| ||| ||||| |||||
3872 LThrProThrThrThrThrProThrThrThrThrThrThrThrThrThrThr 3888
1448 GTGTGAGAGCATGGGTCTCGGACAGCCCGCCAGCATTTACTGGCGCCAG 1497
    ||||| ||| ||||| |||||
3888 hrProThrProThrGlyThrGlnThrProThrThrThrProThrThrThr 3904
1498 GCGCAGTGTGCGCGCTAGTGTACCCCAACTGTACACACATTCACG 1547
    ||||| ||| ||||| |||||

```

3905 ThrThrThrValThrProThrProThrGlyThrGlnThrProTh 3921  
1548 ACACACACACACACACCTCTACACACACTGCA..... 1582  
3921 rThrThrProThrThrThrThrValThrProThrProThrPro 3937  
1583 CGTGGAGGGAGGTGCACACACATCCATTCATTCAGTGTAGAGGAC 1632  
3938 ThrGlyThrGlnThrProThrThrProThrThr.....Th 3951  
1633 CGATCTGCAGAGGACGCGGGGGCCGCCACACACGACAGCTGGAGG 1682  
3951 rThrThrValThrProThrProThrProThr..... 3961  
1683 ATGAGGAGCGAGCTGCAGACGAGGAGGAGCCATGSCGAGAGAA 1732  
3962 .....GlyThrGlnThrProThrThr 3968  
1733 TGGCAGACACCCAGCAGTCTGTGTGAGGATAGCCCTGAGACAC 1782  
3969 ThrProThrThrThrThrThrVal.....ThrProThrProTh 3982  
1783 ACACACAGACACAC.....ACTACTGATGATGATGATGACAC 1823  
3982 rProThrGlyThrGlnThrProThrThrProThrThrThrThr 3999  
1824 ACATGGCGGACAGCTGCTCCGTGAAGGACAGCTAGCAGACAC..... 1867  
3999 hrValThrProThrProThrProThrGlyThrGlnThrProThrThr 4015  
1868 .....CGACATGCACAGATATGCCCTGGGACACAGATTAAGCTGC 1911  
4016 ProThrThrThrThrThrThrValThrProThrProThr.....Pr 4029  
1912 CAATGACGACACAGCAGACAGACATGCCAGAAATACAGAGATGCT 1961  
4029 oThrGlyThrGlnThrProThrThrThrProThrThrThrThrThr 4046  
1962 GCGTGAACATACACAGCAGACCCATGCGAG..... 1993  
4046 alThrProThrProThrProThrGlyThrGlnThrProThrThrPro 4062  
1994 ATGTCTGCTGTGACACACACACACAGGATATGCTGTGAGCGAC 2043  
4063 lThrThrThrThrThrThrValThrProThrProThrProThrGly...Th 4078  
2044 ACAGCTGCAATATGTAATCCGACACACAGCTGCACAGATATGCTGCT 2093  
4078 rGlnThrProThrThrThrProThrThrThrThrThrThrThrPro 4094  
2094 GACACACAGATATGCTGCTGTGACACACACATGACGATATGCTG 2143  
4095 ThrProThrProThrGlyThrGlnThrProThrThrThrProThrThr 4111  
2144 GACACACACACACACAGCTGTGACAGA...TATGCTGTGACAGCGCA 2190  
4111 rThrThrThrValThrProThrProThrProThrGlyThrGlnThrPro 4128  
2191 CACACATGCAATATGCTGCTGTGACACACATCCAGACACAGCTGCAC 2240  
4128 hrThrThrProThrThrThrThrThrValThrProThrProThr... 4143  
2241 AGGCGAGATATGCTGCTGTGACACACGCGATATGCTGTATGTCAC 2290  
4144 .....ProThrGlyTh 4147  
2291 ACACACGACAGATGCTGTGACACACACAGCAGATGATATGCT 2340  
4147 rGlnThrProThrThrThrProThrThrThrThrThrValThrPro 4164  
2341 GTCGAGACACAGCAGCAGATATGCTGCTGTGACACACACAGAGA 2390  
4164 hrProThrProThrGlyThrGln.....ThrProThrTh 4175

2391 TATGCTGCTCAGACACTCAGACAGCTGACAGATATGCTGACACAC 2440  
4175 rThrProThrThrThrThrThrValThrProThrPro..... 4188  
2441 ATGTGACAGATATGCTGTGACATGCACACAGCTGCAGATATGCT 2490  
4189 .....ThrProThrGlyThrGln...ThrGly 4196  
2491 CCGATACACAGCAGC.....ACACATGACA 2519  
4197 ProProThrHisThrSerThrAlaProIleAlaGluThrThrSer 4213  
2520 TATGCTGCTGGGACACACTTCGGACACATGACACACAGCTGAC 2569  
4213 hrProProGluThrSerThrProGlnThrSerThrSerThrSer 4230  
2570 ATATGCTGCTGTGACACAGCAGACTGACTGCTTTGGAGGCTGCT 2619  
4230 rGlu.....ThrGluThrThrLeuLeuSerThrLeuThr 4242  
2620 GTGAGCTGACATGATGCTGCGCTGAGGCTCATAGTATGAGGACTT 2669  
4242 oProAlaIleGluMet.....Thrs 4249  
2670 TCCCTGCTCCACGCTCCTCCCACTGCTGCCG..... 2704  
4249 erThrAlaProProSerThrPro...ThrAlaProThrThrThrSer 4264  
2705 .....CCTGCTGCCGCTCAGTCCCGCCCATCCCGCTCTGTC 2748  
4265 GlyHisThrLeuSerProProProSerThrThrThrSerProThr 4281  
2749 CCTGCTGCTGGGCTATTTTGCACCTGCTGCTGGTGCAGAGTTC 2798  
4281 rProThrThrGly..... 4285  
2799 CCTACTGCTGTGGGCTGGGCTGGGACACAGACCCCA..... 2839  
4286 .....ThrThrThrGlySerSerSerAlaProThrProSer 4297  
2840 .....AGCTGAGAGCTGAGCCCATGCTAGTGTCTCA 2874  
4298 ThrValGlnThrThrThrThrSerAlaThrPro.....ThrProTh 4312  
2875 TCCCACTGATTCCTCCCTGACACAGAGAGGCTGTGATTATTA 2924  
4312 rProLeuSerThrProSerIleIleArgThrThrGlyLeuArgProThr 4329  
2925 TTTAAGAAATGAAGATATATTAATATGATGGA..... 2959  
4329 roSerSerValLeuIleCysCysValLeuAsnAspThrTyTyAlaPro 4345  
2960 .....GGAAGACTGGGT.....TGACG.....GA 2979  
4346 GlyGluGluValTyAsnGlyThrTyGlyAspThrCysTyPheValAs 4362  
2980 CTGTGCTCTTCC.....TGAGGC..... 2998  
4362 ncysSerLeuSerCysThrLeuGluPheTyAsnTrpSerCysProSer 4379  
2999 .....CCGGACCCGCTGCTTTACGCCATGCTGATACACACCCGCTC 3046  
4379 hrProSerProThrProThrProSerThrSerThrPro...ThrProSer 4394  
3047 AGGCGACACACACACCC...CACCCACTGCTGTGAGGCCACATCT 3093  
4395 LysProSerSerThrProSerLysProThrProGlyThrLysProPro 4411  
3094 CTGT 3097  
4411 ucys 4412





```

687 GGTGATGATGATCCAGCGGACCGCTTCAGACCCGCTTCACAGCGACGC 736
      |||||
410 rleuglval.....glutrylprolylsleugllyl.... 421
737 ACCCGCGTAACAGAGCGTGGATTTGGGGGGGACCAAGTCCCTCCAGTGC 786
      |||||
422 ..provalalavaltyrthrtpgluglyasnlnvalasnllernrcys 437
787 AAGTGCAGCAGCAGCAGTGAACCGGTGATCAGTGGCTGAGCGCGTGA 836
      |||||
438 GluValAlaPheLalatyProSerAlaThrIleSerTyrPhearg..... 451
837 GTAAGCGCGCGAGCGCGCCACACTCCACATCGATGGCGCGCCAGA 886
452 .....AspGlyGln.. 454
887 AGTTTGCTGCTGCCACGGGAGAC.....GTGGTGC 921
      |||||
455 .....LeuLeuProSerSerAsnTyrSerAsnIleIleTyrAsn 468
922 CGGCGCGAGCGCTCTACCTCAATAGCTGCTCATCAGCGCGCGCGCA 971
      |||||
469 ThrProSerLaserTyrleugl.....ValThrProAspsergl 482
972 GGAAGATCGCGCATGATCATCTGCTGCGCGCCACACATGGCTACA 1021
      |||||
482 uasnAspPheGlyAsnTyrAsnCystrhrAlaValAsnArgIleGlyGln 499
1022 GCTCCGCGAGCGCTCTCAGCGTGTGTCAGACCCAAACCGCGAGG 1071
      |||||
499 luserleuglupheleleuValGlnAlaAspThrProSerSerProser 515
1072 .....CC 1073
516 lIeAspArgValGluProTyrSerSerThrAlaGlnValGlnPheAspG1 532
1074 ACCGTGCGCTCCGCTGCGCGCGCA...CTAGCGTCCGTGGCGCGTGG 1120
      |||||
532 uprogluAlaThrIleGlyVal..ProIleLeuIleTyrLysAlaGluTTP 548
1121 TCATCGCATCCAGCGCGCGCTGTCTTCATCTCGGACCGCTGCTCC.. 1168
      |||||
549 lYserleuglGluGlnAlaTPrhSerIleTyrLysAlaIleGlyG1 565
1169 .....TGTGCTTCCAGCGCGCGCA 1190
565 uAlaAsnMetGluGlyIleValThrIleMetGlyLeu...LysProGluT 581
1191 GAAGCGCTGACCCCGCGCTGCGCTGCGCGCGCGCGCGCGCGCG 1240
      |||||
581 hrArgTyrAla.....ValArgLeuAlaAlaValAsnGlyLysGly 594
1241 CGGGAGCGCGCGCGCGAGCGCGAGCGAGCAAGACCTCCCTGTTGGCC 1290
      |||||
595 lengllylulleserAlaAlaThrGlnPheLysThrGlnProValArgG1 611
1291 GCCTGACCGCTGCGCTGCTGTTGGCGCTGTGTAGAGAGATGGGT... 1336
      |||||
611 uProSerAlaProLysLeu.....GluGlyGlnMetGlyGluA 624
1337 .....CTCGCGCAG..... 1345
624 spGlyAsnSerIleLysValAsnLeuLeuLysGlnAspArgLysSer 640
1346 CCCCCCAGCATCTACTGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1395
      |||||
641 ProIleArgHisTyrLeuValLysTyrArgAlaLeuAlaSerGluTyrPly 657
1396 CCGAAGCTCTACACAGATCCACACACACACACAGATGATTGGG 1441
657 sprogluIleArgLeuProSerGlySerAspHisValMetLeuLys 672

```

```

seq_name: Swissprot_39:NCAL_HUMAN
seq_documentation_block:
ID NCAL_HUMAN STANDARD: PRT: 848 AA.
AC P13591; Q16180; Q15829;
DE 01-JAN-1990 (Rel. 13, Created)
DE 15-JUL-1998 (Rel. 36, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEURAL CELL ADHESION MOLECULE, 140 KDA ISOFORM PRECURSOR (N-CAM 140)
DE (NCAM-140) (CD56 ANTIGEN).
DE NCAM1 OR NCAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94356433; PubMed=8075973;
RA Saito S., Tani Y., Tachibana I., Hayashi S., Kishimoto T., Kawase I.;
RT "Complementary DNA sequence encoding the major neural cell adhesion
RT molecule isoform in a human small cell lung cancer cell line.";
RL Lung Cancer 10:307-318(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91250739; PubMed=1710251;
RA Lanter L.L., Chang C., Azuma M., Ruitenberg J.J., Hemperly J.J.,
RA Phillips J.H.;
RT "Molecular and functional analysis of human natural killer cell-
RT associated neural cell adhesion molecule (N-CAM/CD56).";
RL J. Immunol. 146:4421-4426(1991).
RN [3]
RP SEQUENCE OF 491-848 FROM N.A.
RX MEDLINE=87301755; PubMed=2887295;
RA Dickson G., Gower H.J., Barton C.H., Prentice H.M., Elsom V.L.,
RA Moore S.E., Cox R.D., Quinn C., Putt W., Walsh F.S.;
RT "Human muscle neural cell adhesion molecule (N-CAM): Identification
RT of a muscle-specific sequence in the extracellular domain.";
RL Cell 50:1119-1130(1987).
CC -!- FUNCTION: THIS PROTEIN IS A CELL ADHESION MOLECULE INVOLVED IN
CC NEURON-NEURON ADHESION, NEURITE FASCICULATION, OUTGROWTH OF
CC NEURITES, ETC.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- ALTERNATIVE PRODUCTS: THE DIFFERENT TISSUE-SPECIFIC FORMS OF
CC N-CAM ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -!- DATABASE: NAME=PROM; NOTE=CD guide CD56 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd56.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL, S71824; AAB31836.1; -
DR EMBL, U63041; AAB04558.1; -
DR EMBL, M17410; AAB59913.1; -
DR PIR, B26883; B26883.
DR MIM, 116930; -
DR InterPro, IPR001777; -
DR InterPro, IPR003006; -
DR Pfam, PF00041; fn3; 2.
DR Pfam, PF00047; igf; 5.
DR Immunoglobulin domain; Cell adhesion; Glycoprotein; Repeat; Signal;
KW Transmembrane; Alternative splicing.
FT SIGNAL 1 19
FT CHAIN 20 848 NEURAL CELL ADHESION MOLECULE, 140 KDA
FT DOMAIN 20 708 ISOFORM.
FT TRANSMEM 709 729 EXTRACELLULAR (POTENTIAL).
FT

```



```

1345 GCGCCCGAG 1353
|||||:::
604 Alaprollys 606

seq_name: SwissProt_39:EG15_CAEEL
seq_documentation_block:
ID EG15_CAEEL STANDARD: PRT: 1040 AA.
AC Q10656:EG15_CAEEL
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE MYOBLAST GROWTH FACTOR RECEPTOR ELG-15 PRECURSOR (EC 2.7.1.112).
GN EG1-15
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=96069862; PubMed=7585964;
RA Devore D.L., Horvitz H.R., Stern M.J.;
RT "An FGF receptor signaling pathway is required for the normal cell
RL Cell 83:611-620(1995).
CC -1- FUNCTION: RECEPTOR FOR AN EXTRACELLULAR SIGNAL REQUIRED FOR
CC THE GUIDANCE OF SEX MYOBLAST MIGRATION DURING GONAD DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isdb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
CC -----
DR EMBL: U39761; AAC46934.1; -
DR HSP: P11362; 1FGI.
DR InterPro: IPR000719; -
DR InterPro: IPR001245; -
DR InterPro: IPR003006; -
DR Pfam: PF00047; Ig_3.
DR Pfam: PF00069; PKinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR Receptor: Glycoprotein; Tyrosine-protein kinase; ATP-binding;
KW Transferrase; Phosphorylation; Transmembrane; Signal;
KW Immunoglobulin domain.
KW SIGNAL
FT CHAIN 1 19 POTENTIAL.
FT DOMAIN 20 1040 MYOBLAST GROWTH FACTOR RECEPTOR ELG-15.
FT TRANSMEM 526 549 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 550 1040 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 48 116 IG-LIKE DOMAIN.
FT DOMAIN 307 374 IG-LIKE DOMAIN.
FT DOMAIN 407 492 IG-LIKE DOMAIN.
FT DOMAIN 640 931 PROTEIN KINASE.
FT DOMAIN 646 654 ATP (BY SIMILARITY).
FT BINDING 672 672 ATP (BY SIMILARITY).
FT ACT_SITE 797 797 BY SIMILARITY.
FT MOD_RES 828 828 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT DISULFID 314 367 POTENTIAL.
FT DISULFID 414 485 POTENTIAL.
FT CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 407 407 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 449 449 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 474 474 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 497 497 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT MUTAGEN 680 680 E-X: IN N1775; LOSS OF ACTIVITY.
FT MUTAGEN 714 714 P-DL: IN N1783; LOSS OF ACTIVITY.
FT MUTAGEN 930 930 W->*: IN N14775; LOSS OF ACTIVITY.
SQ SEQUENCE 1040 AA; 118956 MW; 8CA227195D8CD69 CRC64;

alignment_scores:
Quality: 256.50 Length: 419
Ratio: 1.386 Gaps: 12
Percent Similarity: 44.153 Percent Identity: 20.764

alignment_block:
US-09-598-042a-2 x EG15_CAEEL
Align seg 1/1 to: EG15_CAEEL from: 1 to: 1040

232 CTGACGTCAGACGAGTCGAGCGGAGATCGCGGCTGACGTGCA 281
||| ::|||::: ::||| ::|||:::
94 LEUNHLEUHLNLSANLALTYGLINLESERARGASPHETHVALGLUVAL 110
::: ::|||::: ::|||::: ::|||:::
282 GCGCCACCAACGGCTTGCGGACCGCTGACGCTCACTACCCGCTGCTGC 331
::: ::|||::: ::|||::: ::|||:::
110 GVALHLSANLALTYGLINLESERARGASPHETHVALGLUVAL 127
332 TGGAT. 336
127 LESPHECYASPTYPHELEUPHEPROASPLEHSHLSLEUANL 143
336 336
144 PROMETGLUCYSVALCYSLPUTRPLTYRASNLYGLUALALYSARGSE 160
337 GACATTAGC. 345
160 RASPVASLNTYRALALALVALTHRGUVALCYSELYSTYVALAS 177
345 345
177 TERARGMETLEASNARGALALARGLYSPROLEUPROMETILEPROCYSPHE 193
345 345
194 GLYASPHISCYLSYSGLUPEASPTHTHTHPROVALSERASPHGLYLE 210
346 CGAGGAAGAGAGAGCGCTGGGCC. 369
210 UPTRGGLYLYSPROGLIUSPASPROLEUVALLYSARGVALLEULYS 227
370 370
227 YASASPVALLILEVALPROVALHISASPER. 239
394 GACCCCGCCGACGACGAGTGGGACGACCGCGCTC. 429
240 SERPROSERGLU. 253
430 430
253 ULYSGLUASLYSGLUASPELUGLIUSPTYSERVALSERGLINPROV 270
438 438
270 ALALAPROASPLALAGLYLEUTHRGLEUASNLLETHRALAGLUPTRO 286
439 TCAGATGAGGCGCGGCTGATC. 464

```

```
287  ProTyrPheLysSerAsnAspAsnIleValLeuPheAsnGluThrHisAl 303
      |||  ...  |||
465  AGCGCCGCGTGGTACGTCGCGGCTCAAGTCGCGCCAGCGGGCACC 514
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
303  AlenPrAlaIcLgIaThrLeuLysLeuAsnGlySarGlyAlaLysGlyTyr 320
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
515  CTGGCCGCGCATCAGCGATGAAAGCAGCAGCCGCTTGACCGCCCA 564
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
320  roGluProGlnIleIleThrTyrLysAsnGlyLysMetLeuLysSer 336
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
565  GAGCGCGCT.....GAGCCGAGAAAGATGAGACACTAG 602
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
337  SerIleAsnGserGlyGlyTyrGluPheLysPheAsnArgTyrPserLeuG 353
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
603  CCGTGAAGAACCTGGCGCGGAGAGCAGCGGCAATACACTGCGCGCTGT 652
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
353  uValGluAspAlaValAlaAspSerGlyGluPheHisCysGluAlaL 370
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
653  CGAACCGCGCGCGCGCATCAACCCGACCTACAAAGTGATGTATCCAG 702
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
370  euAsnLysValGlySerAlaLysTyrThrHisValIleIleValAsn 386
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
703  CGGACCGCGCTCCAGCCGCTGCTACAGCAGCGACCCCGTGAACAGCAG 752
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
387  ArgMetArgArgProPheIleIleValProAsnIleLeuAlaAsnGlnSe 403
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
753  GGTGGACTTCGGGGGAGCAGCAGCTCTCCAGTGCAGAGTGCGCAGCAGC 802
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
403  rValAsnIleAsnAspThrIleAlaThrPheHisCysLysValAlaSerAsp 420
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
803  TGAACCGCGGTATCGAGTGGCTGAACCGCGGTGAGTAC..GGCGCCGAG 849
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
420  euLysProHisIleIleThrValArgLysLysLysLysLysLysLysTyr 436
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
850  GCGCGCCGACATCCACCATCATGATGGCGGCGCAGAGTTGGTGGTCT 899
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
437  SerTyrTyrAsnAsnSerAlaGlnGluTyrMetPheAsnTyrThrGluMe 453
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
900  GCCCAGCGGTGAC.....GTGTGGTCCGGCGCGCGCGCTCTACCC 940
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
453  tAspThrPheAsnLysAlaHisValHisHisValGlyAspLysLysTyr 468
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
941  TCAATAGCTGCTCATCAGCCGCGCGCGCGCAGAGATCGCGCATGTAC 990
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
469  ....ThrLeuThrIlePheAsnValSerLeuAspAspGlnGlyLysTyr 483
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
991  ATTCGCTTGGCGCAACACCATGAGCTTCCGCGCGCGCTTCT 1040
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
484  AlAcysLeuSerGlyAsnSerLeuGlyMetSerMetAlaAsnAlaThrLe 500
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1041  CACCGGTG 1047
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
500  uThrVal 502
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

seq_name: SwissProt_39:CONT_MOUSE

seq_documentation_block:
  CONT_MOUSE STANDARD; PRT; 1020 AA.
  AC P12960;
  DT 01-JAN-1990 (Rel. 13, Created)
  DT 01-JAN-1990 (Rel. 13, Last sequence update)
  DT 15-JUL-1999 (Rel. 38, Last annotation update)
  DE CONTRACTIN PRECURSOR (NEURAL CELL SURFACE PROTEIN F3).
  GN CNTN1.
  OS Mus musculus (Mouse).
  OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  OX NCBT_TaxID=10090;
  RN [1]
  RP SEQUENCE FROM N.A.
  RC STRAIN=C57BL/6; TISSUE=Brain;
  RX MEDLINE=89340657; PubMed=2474555;
```

```
RA Gennarini G., Cibelli G., Rougon G., Mattei M.-G., Goridis C.;
RT "The mouse neuronal cell surface protein F3: a phosphatidylinositol-
RT anchored member of the immunoglobulin superfamily related to chicken
RT contactin."
RL J. Cell Biol. 109:775-788(1989).
CC -1- FUNCTION: MEDIATES CELL SURFACE INTERACTIONS DURING NERVOUS
CC SYSTEM DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
CC -1- MISCELLANEOUS: F3 SHARES WITH L1, N-CAM, MAG, AND OTHER CELL
CC ADHESION MOLECULES FROM NERVOUS TISSUE THE L2/HNK-1 CARBOHYDRATE
CC EPITOPE.
CC -1- SIMILARITY: CONTAINS 4 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 6 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X14943; CAA33075.1; -.
CC DR HSSP: S05944; S05944.
CC DR MGD; MG1:105980; 1BCU.
CC DR InterPro; IPR001777; -.
CC DR InterPro; IPR003006; -.
CC DR Pfam; PF00041; fn3; 4.
CC DR Pfam; PF00047; lg; 6.
CC KW Immunoglobulin domain; Glycoprotein; Signal; GPI-anchor;
CC Cell adhesion; Repeat.
CC FT SIGNAL 1 20
CC FT CHAIN 1 1020
CC FT PROPEP 58 121
CC FT DOMAIN 151 218
CC FT DOMAIN 256 317
CC FT DOMAIN 345 398
CC FT DOMAIN 429 491
CC FT DOMAIN 519 592
CC FT DOMAIN 604 611
CC FT DOMAIN 611 712
CC FT DOMAIN 713 814
CC FT DOMAIN 815 910
CC FT DOMAIN 911 1006
CC FT DISULFID 65 114
CC FT DISULFID 158 211
CC FT DISULFID 263 310
CC FT DISULFID 352 391
CC FT DISULFID 436 484
CC FT DISULFID 526 585
CC FT CARBOHYD 208 208
CC FT CARBOHYD 258 258
CC FT CARBOHYD 338 338
CC FT CARBOHYD 457 457
CC FT CARBOHYD 473 473
CC FT CARBOHYD 494 494
CC FT CARBOHYD 521 521
CC FT CARBOHYD 593 593
CC FT CARBOHYD 935 935
CC SQ SEQUENCE 1020 AA; 113388 MW; 9DCDAA40EAA4C8C7 CMC64;

alignment_scores:
  Quality: 255.50 Length: 397
  Ratio: 1.352 Gaps: 16
  Percent Similarity: 47.607 Percent Identity: 23.929

alignment_block:
  us-09-598-042a-2 x CONT_MOUSE
  Align seg 1/1 to: CONT_MOUSE from: 1 to: 1020
```

```

58 ATCCAGTATGTCTCCGGGT.....CGACGATTTCGTCGACAGT 101
   |||:|||||:|||||:
325 ILfTYrValGlInAlaPheProGluTrValGlInHisIleAsnAspThrGl 341
102 GGGCCGGCTGGGGCGACATATGGGCTGCAGTCCGACAGGAGGAGGAGC 151
   |||:|||||:|||||:
341 uValAspIleGlySerAspLeuTrTPProCysIleAlaThrGlyLysP 358
152 CGCCGGCGCTGACCATGTGACCAAGAGATGGCCGACCATCCACAGCGG 201
   |||:|||||:|||||:
358 rIleProThrIleArgTrpLeuLysAsnGlyTrSerTrpHisLysGly 374
202 TGGAGCCGCTCCGGCTGCTCCGCGACGGGCTAAGGTGAGACAGGTGA 251
   |||:|||||:|||||:
375 .....GluLeuAlaArgLeuTrpAspValTh 382
252 GCGGAGAGATCGGGCGGTACGTACGTGACAGGCCACACAGCGCTCGCA 301
   |||:|||||:|||||:
382 rPheGluAsnAlaGlyMetTrpGlnCysIleAlaGluAsnAlaTrpGlyS 399
302 GCCTGACCGCTCAACTACACCTCGTCTGTGATGACATTAGCCCAAGG 351
   |||:|||||:|||||:
399 eTrIleTrpAlaAsnAlaGluLeuLysIleLeu..... 409
352 AAGAGAGCGCTGGGGCGCCGACAGCTCTCTGGGGGTCAAGAGACCCCG 401
   |||:|||||:|||||:
410 .....AlaLeuAlaProThr..... 414
402 CAGCCAGCAGTGGGCGACAGCGCGCTTCACACAGCCCTCCAGATGAGC 451
   |||:|||||:|||||:
415 .....PheGluMetAsnProMetLysL 422
452 GCGGGGTGATGCGACGGCCCGTGGTACCTCCGCGGCTCAAGTGGTG 501
   |||:|||||:|||||:
422 ySLysIleLeuAlaAlaLys...GlyGlyArgValIleIleGlyCysLys 437
502 GCGCGGGGCGCCTCGGCGGCATACAGTGTGTAAGAGCAGCAGCGC 551
   |||:|||||:|||||:
438 ProLysAlaAlaProLysProLysPheSerTrpSerLysGlyThrGluTr 454
552 CTTCAGCGCGCCAGAGCGCGCTGACCCAGAGAGAGAGAGAGTGG..... 594
   |||:|||||:|||||:
454 pLeuValAsn.....SerSerArgIleLeuIleTrpLysAspG 467
595 ..ACACGTAGCGCTGAAGACCTGCGCGCGGAGAGCAGCGCAATATACC 642
   |||:|||||:|||||:
467 ySerLeuGluIleAsnAsnIleThrArgAsnAspGlyGlyIleTrpThr 483
643 TGCCGCGGTGACAGCGCGCGGCGCATCAACGCCACCTCAAGTGA 692
   |||:|||||:|||||:
484 CysPheAlaGluAsnAsnArgGlyLysAlaAsnSerThrGlyThrLeuVa 500
693 TGTGATCCAGCGGACCGCTTCAGAGCCGCTGTCACAGGACGACCCCG 742
   |||:|||||:|||||:
500 ILleThrAsnProThrArgIle.....IleLeuAla.....ProI 512
743 TGACACAGCGGTGACTTCGGGGGACACAGTCTCTCCAGTGAAGTG 792
   |||:|||||:|||||:
512 IeAsnAlaAspIleThrValGlyLysAlaAlaThrMetGlnCysAlaAla 528
793 CGACGACGATGAAGCGGTGATCCAG.....TGGCTGAAGCG 830
   |||:|||||:|||||:
529 SerPheAsp.....ProAlaLeuAspLeuThrPheValTrpSerPheAs 543
831 C.....GTGAGTAGCGCGCGCGGAGCGCCGAC..... 858
   |||:|||||:|||||:
543 nGlyTrValIleAspPheAsnLysGluIleThrHisIleHisTrpGlna 560
859 ..AACTCCACCATGATGTGGCGCGCGAGAGTTTGCTGCTGCGCCAG 906
   |||:|||||:|||||:
560 rGAsnPheMetLeuAspAlaAsnGlyLys..... 569

```

```

907 GGTGACGTGTGTCGGCGCGCCGACGCGCTCTCACTCAATTAAGTCTCAT 956
   |||:|||||:
570 .....LeuLeuI 572
957 CACCCGTGCGCCGACGAGAGATGCGGCATGATCATCTGCTGGCGCA 1006
   |||:|||||:
572 eArgAsnAlaGlnLeuLysHisAlaGlyArgTrpThrCysThrAlaGlnT 589
1007 ACACCATGGGCTACAGCTTCGCCAGCGCTTCCTCAGCGCTGTCGCAGAC 1056
   |||:|||||:
589 hTrIleValAspAsnSerSerAlaSerAlaAspLeuValValArgGlyPro 605
1057 CCAAAACGCGACGAGCGCACCTGTGCGCTCTCTGCTGTCGCGCATAGCCT 1106
   |||:|||||:
606 ProGlyProProGlyGlyLeuArgGlyIleGluAspIleArgAlaThrSerVa 622
1107 GCGGTGCGCGCTGCTCATCGCATCC..... 1132
   |||:|||||:
622 lAlaLeu.ThrTrpSerArgGlySerAspAsnHisSerProIleSerLys 638
1133 .....CAGCCGCGCTGCTTCATCCTGG 1156
639 TyrThrIleGlnThrLysThrIleLeuSerAspAspTrp 651
seq_name: SwissProt_39:NTC1_RAT
seq_documentation_block:
ID NTC1_RAT STANDARD: PRT: 2531 AA.
AC 007008:
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1999 (Rel. 38, last sequence update)
DT 01-OCT-2000 (Rel. 40, last annotation update)
DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1 PRECURSOR.
GN NOTCH1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Schwann cell;
RX MEDLINE=92111383; PubMed=1764995;
RA Weinmaster G., Roberts V.J., Lemke G.;
RT "A homolog of Drosophila Notch expressed during mammalian
RT development 113:199-205(1991).
RL development 113:199-205(1991).
RN [2]
RP REVISIONS TO 1652-1653.
RA Weinmaster G.;
RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: REQUIRED FOR THE CORRECT DIFFERENTIATION OF A NUMBER
CC OF TISSUES.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGHEST LEVELS OCCUR BETWEEN
CC DAYS 12 AND 14 AND DECREASE RAPIDLY TO MUCH LOWER LEVELS IN THE
CC ADULT.
CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X57405; CAA40667.1; -.
CC DR HSSP: P00740; IIXA.
CC DR InterPro: IPR000152; -.
CC DR InterPro: IPR000561; -.

```









CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 DR EMBL: X12875; CA31368.1; -.  
 DR PIR: S05479; S05479.  
 DR HSSP: P20241; ICEB.  
 DR MGD: MGI:96721; L1cam.  
 DR InterPro: IPR001777; -.  
 DR InterPro: IPR003006; -.  
 DR Pfam: PF00041; fn3; 4.  
 DR Pfam: PF00047; 1g; 6.  
 DR PRINTS: PR00014; FNTYPE11.  
 DR Cell adhesion; Glycoprotein; Transmembrane; Repeat; Brain;  
 DR Immunoglobulin domain; Signal; Alternative splicing.  
 DR SIGNAL  
 FT CHAIN 1 19  
 FT DOMAIN 20 1260 NEURAL CELL ADHESION MOLECULE L1.  
 FT TRANSMEM 20 1123 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 1124 1146 POTENTIAL.  
 FT DOMAIN 1147 1260 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 50 120 IG-LIKE C2-TYPE DOMAIN.  
 FT DOMAIN 150 215 IG-LIKE C2-TYPE DOMAIN.  
 FT DOMAIN 256 318 IG-LIKE C2-TYPE DOMAIN.  
 FT DOMAIN 346 410 IG-LIKE C2-TYPE DOMAIN.  
 FT DOMAIN 440 503 IG-LIKE C2-TYPE DOMAIN.  
 FT DOMAIN 531 599 FIBRONECTIN TYPE-III.  
 FT DOMAIN 827 896 FIBRONECTIN TYPE-III.  
 FT DOMAIN 932 994 FIBRONECTIN TYPE-III.  
 FT DOMAIN 1032 1094 FIBRONECTIN TYPE-III.  
 FT SITE 553 555 CELL ATTACHMENT SITE (POTENTIAL).  
 FT SITE 562 564 CELL ATTACHMENT SITE (POTENTIAL).  
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 432 432 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 478 478 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 504 504 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 587 587 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 725 725 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 776 776 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 824 824 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 848 848 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 875 875 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 968 968 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 978 978 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1022 1022 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1030 1030 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1073 1073 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1107 1107 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 1180 1183 MISSING (IN SHORT ISOFORM)  
 FT  
 SQ SEQUENCE 1260 AA; 140968 MW; 22B857001CB2A538 CRC64;  
 (BY SIMILARITY).

## alignment\_scores:

Quality: 252.50 Length: 829  
 Ratio: 0.726 Gaps: 39  
 Percent Similarity: 41.978 Percent Identity: 22.557

## alignment\_block:

US-09-598-042a-2 x CAML\_MOUSE ..

Align seg 1/1 to: CAML\_MOUSE from: 1 to: 1260

```

73 TTCCCGGCTGACGATTTCTCGGCAAGTGGCGGCGGCGACTAT 122
   |||||  ::  |||  |||||  |||:::
243 PheProThrsAsnSerSerArgLeuValAlaLeuGlnGlyInsSer 259
   |||||  ::  |||||  |||||  |||||
123 GCGGCTGACGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 172
   |||||  ::  |||||  |||||  |||||
259 uIleLeuGlnCysIleAlaGlnGlyPheProThrProThrIleLysTrpL 276

```

```

173 CCAAGGATGGCCGACCAATCCACAGCGGCTGGAGCGCTTCGCGCTG 222
   ::::  ::::  ::::  |||||
276 euHisProSerAspProMetProThrAsp.....ArgValIle 288
   ::::  ::::  ::::  |||||
223 CCGCAGGG.....CTGAAGTGAACACAGTGAAGCGGAGGA 260
   |||||  ::::  ::::  |||||
289 TyrGlnAsnHisAsnLysThrLeuGlnLeuLeuAsnValGlyGlnLys 305
   |||||  ::::  ::::  |||||
261 TGCGGCGGTGATGTCGCAAGGCCCAACAGGCGCTCGGCGGCGTGA 310
   |||||  ::::  ::::  |||||
305 pAspGlyGlnTyrThrCysLeuAlaGlnSerLeuGlySerAlaArgH 322
   |||||  ::::  ::::  |||||
311 TCAACTACACCCCTGCTGCTGGATGACATTAGCCAGGAGAGAGAGC 360
   |||||  ::::  ::::  |||||
322 IsAlaTyrTyrValThrVal..... 328
   |||||  ::::  ::::  |||||
361 CTGGGCGCCGACAGCTCTCTGGGGGTCAAGAGAGCCGCCAGCAGCA 410
   |||||  ::::  ::::  |||||
329 .....GluAlaIleProty 333
   |||||  ::::  ::::  |||||
411 GTGGGACGACGCGGCTTCACACAGCCCTCCAGATGAGGCGCGGTGA 460
   |||||  ::::  ::::  |||||
333 rTrp..... 335
   |||||  ::::  ::::  |||||
461 TCGCAGGCGCC.....GTGGTAGCTCCGTCGCGCTC 492
   ::::  ::::  ::::  |||||
335 euGlnLysProGlnSerHisLeuTyrGlyProGlyGlnThrAlaArgLeu 351
   |||||  ::::  ::::  |||||
493 AAGTGGCTGCGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 542
   |||||  ::::  ::::  |||||
352 AspCysGlnValGlnIleTyrProGlnProGlnIleThrTyrArgLys 368
   |||||  ::::  ::::  |||||
543 CGACGAGCGCTTG...ACGGCCGACAGGCGGCTGAGCCAGGAAGA 589
   ::::  ::::  ::::  |||||
368 nGlyMetSerMetGlnThrValAsnLysAspGlnLysTyrArgIleGln 385
   |||||  ::::  ::::  |||||
590 AGTGAAGTGAAGCTGTAAGAACCTCGGCGGCGGAGAGAGCGGAATAC 639
   ::::  ::::  ::::  |||||
385 IIndLysLeuIleLeuSerAsnValGlnProThrAspThrMetValThr 401
   |||||  ::::  ::::  |||||
640 ACCTGCGCGGCTGTCGAACCGCGGCGGCGGCGGCGGCGGCGGCGG 689
   |||||  ::::  ::::  |||||
402 GlnCysGlnAlaArgAsnGlnHisGlyLeuLeuAlaAsnAlaTyrIle 418
   |||||  ::::  ::::  |||||
690 GGATGGATTCACAGCGGCGGCTTCACAGCGGCGGCTGTCACAGCGCAC 739
   ::::  ::::  ::::  |||||
418 eTyValAlaGlnLeuProAlaArg.....IleLeuThrLysAspAsn 433
   |||||  ::::  ::::  |||||
740 CCGTGAACAGCAGGTGACTTCGGGCGGAGCACGTCCTTCAGTGCAG 789
   ::::  ::::  ::::  |||||
433 IIndThrMetAlaValGln...GlySerThrAlaTyrLeuLeuCysLys 448
   |||||  ::::  ::::  |||||
790 GTGCGGACGACGCTGAAGCGGCTGATCCAGTGGCTGAAGCGCGTGA 839
   ::::  ::::  ::::  |||||
449 AlaPheGlyAlaProValProSerValGlnTrpLeuAspLeu..... 462
   |||||  ::::  ::::  |||||
840 CCGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 889
   |||||  ::::  ::::  |||||
463 .....GlnGly.....ThrThrValLeuGlnAspGlnArg 473
   |||||  ::::  ::::  |||||
890 TTGTGTGCTGCTCCACGCGGTGACGTGTGTCGCGGCGGCGGCTTCAC 939
   |||||  ::::  ::::  |||||
473 hepHe.....Protyr 476
   |||||  ::::  ::::  |||||
940 CTCGAATAGTGTCTCATCACCGCTGCGCGGCGGCGGCGGCGGCGG 986
   |||||  ::::  ::::  |||||
477 AlaAsnGlyThrLeuSerIleArgAspLeuGlnAlaAsnAspThrLys 493
   |||||  ::::  ::::  |||||
987 GTACATCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1036
   |||||  ::::  ::::  |||||
493 gTyrPheCysGlnAlaAlaAsnAspLysAsn..... 504

```

1037 TCCTACCGCTGTGCGACAGCCAAACCGCGACGCTGTGGCTCC 1086  
:::|||||::: |||:::  
505 ..ValThrIleLeuAlaIasnLeuGln.....ValIysGlu 515  
1087 TCGTCTCGGCGCCACTAGCCTCCGTCGCCGTGTCATCGGCATCCAGC 1136  
:::|||||::: |||:::  
516 AlaThrGlnIleThrGln.....GlyProAsr 524  
1137 CGGCGCTGTCTTCATCTGGGACCCCTGCTTGGCTTGCCAGGCC 1186  
:::|||||::: |||:::  
524 gSerAlaIleGlnIuysIysGlyAlaArgValThrPheThrCysGlnAlas 541  
1187 AGAAGAAGCCGTGCACCCCGCGCCTGCCCTCCCTCCCTGGGACGCC 1236  
:::|||||::: |||:::  
541 erPheAspProSerLeuGlnAlaSerIleThr.....TyrArg 553  
1237 CCGCGCGGAGCGCCCGCGCGACGCGAGACAGAC..... 1275  
|||::: |||:::  
554 GlyAspGlyAsrPleuGlnGluArgGlyAspSerAspIysTyrPheI 570  
1276 .....CTTCCCTGCTGGCGCGCCCTCAAGCCTGGCC 1306  
::: |||::: |||:::  
570 eGluAspGlyIysLeuValIleGlnSerLeuAspTyrSerAspGlnGly 586  
1307 CTGGTGTGGGCTGTGTGAGAGCATGGTCTCCGCGACGCCGCCAGAC 1356  
::: |||::: |||:::  
587 ..AsnTyrSerCysValAlaIasrThrGlnLeuAspGluValGlySer.. 601  
1357 TTACTGGGCGGAGCCAGTGTGCTGCTAAGTGTACCCCAACTCTA 1406  
::: |||::: |||:::  
602 .....ArgAlaGlnLeuLeuValValGlySer..ProGlyProVal 614  
1407 CACAGACATCCACAC.....ACACACACATGTATGGCG 1441  
|||::: |||:::  
615 .....ProHISLeuGlnLeuSerAspArgHISLeuLysGlnSe 628  
1442 CCGGCTG.....TGTGAGAGCATGGTCTCG... 1469  
::: |||::: |||:::  
628 rGlnValHISLeuSerTyrSerProAlaGlnAspHISAsnSerProIleG 645  
1470 .....GCAGCCCGCCAGCAGCTTA 1487  
|||::: |||:::  
645 IuIysTyrAspIleGluPheGluAspIysGluMetAlaProGluIysTyr 661  
1488 CTGGGCGGCGCCAGTGTGCTGCTAAGTGTACCCCAACTCTACAC 1537  
::: |||::: |||:::  
662 PheSerLeuGlnGlyLysAlProGlyAsnGln..ThrSerThrLeuLys 677  
1538 AGACATCCACACACACACACACTCTACACACTCAGCTGAG 1587  
|||::: |||:::  
678 LeuSerProTyrValHISThrPheArgValThrAlaIleAsnLysTyr 694  
1588 AGGCAAGGTCACAGCATCCACTATCATGTGTAGAGCGACCC... 1633  
||| ||| |||::: |||:::  
694 rGlyProGlyIuProSer...ProValSerGlnSerValValThrProG 710  
1634 .....GTATCTGCAGAGGC..... 1648  
||| ||| |||::: |||:::  
710 IuAlaIaProGluIuLysAsnProValAspValArgGlyGlnGlyAsnGlu 726  
1649 .....ACGGGGGGCGCGCCAGACAGGACAGCTGGGA 1680  
||| ||| ||| |||::: |||:::  
727 ThrAsnAsnMetValIleThrTyrPyrProLeuArgTyrPheAspTyrPas 743  
1681 GGATGGAGAGGAGCTGACAGACAGAGGAGGAGCCATGCGGAGG 1730  
::: |||::: |||:::  
743 nAlaProGlnIleGlnTyrArgValGlnTyrArgProGlnGlyLysGlnG 760  
1731 AATGGCCAGACCCAGGCACTGTGTGTGTAGAGCATAGCCCTGGACAC 1780  
|||::: |||::: |||::: |||:::  
760 IuThrTyrArgLysGlnThrValSerAspProPheLeuValValSerAsn 776  
1781 ACACACACAGACACACACTACTGATGATGATGACACACATGCG 1830

||| |||::: |||:::  
777 ThrSerThr.....PheValProTyrGlnIleIely 786  
1831 CGCACACGTGCTC.....CCTGAGGACACACGTACGC 1862  
::: |||::: |||:::  
786 sValGlnAlaValAsnAsnGlnGlyLysGlyProGluIuProGlnValThrI 803  
1863 ACACACGCACATGCACATATGCGCC..... 1890  
::: |||::: |||:::  
803 IeGlyTyrSerGlyLysAspTyrProGlnValSerProGluLeuLysP 819  
1891 .....TGGGCACACAGATA 1904  
820 IleThrIlePheAsnSerSerThrValLeuValArgTyrPyrProValAs 836  
1905 AGCTGCCCAAAATGCACGACGACAGACATGCGAGAACATAC... 1950  
|||::: |||::: |||:::  
836 PleuAlaGlnValLysGlnHISLeuLysGlyTyrAsnValThrTyrTyr 853  
1951 ..AAGACATGCTGCTGACATACA.....CAGGCACACCATGGCGA 1992  
||| |||::: |||:::  
853 rPlyGlySerGlnArgLysHISSerLysArgHISLeHISLysSerHIS 869  
1993 GATGTCGTGCGCGACACACA..... 2013  
|||::: |||::: |||:::  
870 IleValValProAlaAsnThrThrSerAlaIleLeuSerGlyLeuArgPr 886  
2014 .....CACACACAGATATGCTGTC... 2034  
|||::: |||::: |||:::  
886 oTyrSerSerTyrHISValGlnValGlnAlaPheAsnGlnIysGlyLeuG 903  
2035 .....TGCAGCGACACAGCTGCAGATATGATCGGAGAC 2070  
|||::: |||::: |||:::  
903 IyrProAlaSerGlnTyrThrPheSerThrProGlnGlyValProGlnHIS 919  
2071 ACAGCTGCA.....CAGATATGCTGCTGACACACAGATATGCTG... 2112  
||| ||| ||| |||::: |||:::  
920 ProGlnAlaLeuHISLeuGlnCysGlnSerAspTyrSerLeuLeuHIS 936  
2113 .....CCTTGACACACACATGCACAGGATATGCTGCTGACACACACA 2155  
||| ||| |||::: |||:::  
936 strPrlnProProLeuSerHISAsnGlyValLeuThrGlnGlyTyr... 950  
2156 CACACGTGTGCACAGATATGCTGTGAGACAC 2187  
::: |||::: |||:::  
951 .....LeuLeuSerTyrHIS 955

10/1/01

---



```

|||||
132 rserGIyGIyGIyGIyAspProAlaSerGIyGlnTrpAlaArgProArgp 149
428 TCACAGAGCCCTCCAGATGAGGAGCGCGGTGATCCGACGCGCGGTGGCT 477
149 heThrGlnProSerLysMetArgArgArgValIleAlaArgProValGIy 165
478 AGCTCCGTGGGCTCAAGTGCCTGGCCAGCGGCGACCCCTGGCCGACAT 527
166 SerSerValArgLeuLysCysValAlaSerGIyHisProArgProAspIle 182
528 CACGTGATGAGAGGACGACGCGCTTGACGCGCCCAAGAGCGCGCTGAGC 577
182 eThrTrpMetLysAspAspGlnAlaLeuThrArgProGlnAlaAlaGlnP 199
578 CCGAGGAAGAAGATGGACACCTGAGCTGGAAGAACTGCGCGCGGAGAGAC 627
199 roArgLysLysLysTrpThrLeuSerLeuLysAsnLeuArgProGlnAsp 215
628 AGCGGAAATACACTGCGCGCTGTGGAACCGCGGCGCGCTCAACGCG 677
216 SerGIyLysTrpThrCysArgValSerAsnArgAlaGlyAlaIleAsnAl 232
678 CACCTACAGGTGATGTGATCCAGCGGACCGCTCCAGCGCGCTGCTCA 727
232 aThrTrpLysValAspValIleGlnArgThrArgSerLysProValLeuTr 249
728 CAGGACGACGACCGCGGTGACACGACGCGTGCAGCGGCGGAGACAGCTC 777
249 hTrIYThrHisProValAsnThrThrValAspPheGIyGIyTrpThrSer 265
778 TTCACATGCAGAGTGGCGAGCGACGTGAAGCGGTGATCCAGTGGTGAA 827
266 PheGlnCysLysValArgSerAspValLysProValIleGlnTrpLeuL 282
828 GCGGTGAGATGACGCGCGGAGGCGCGCAACATCCACCATGATGTGG 877
282 sarValGIyTrpGIyAlaGIyLysArgHisAsnSerThrIleAspValG 299
878 GCGGCGAAGATTGTGGTGTGCTGCCACGAGGTGATCGTGGTGGCGCC 927
299 LysGIyLysLysPheValIleValLeuProThrGIyAspValTrpSerArgPro 315
928 GAGGCGTCCACTCAATAGCTGTATCACCCTGACCGCGCGGACGAGCGA 977
316 AspGIySerTrpLeuAsnLysLeuLeuIleThrArgAlaArgGlnAspAs 332
978 TGGGCGATGTATCATCTGCTTGGCGCAACACCATGGGCTACAGCTTC 1027
332 palagIyMetTrpIleCysLeuGIyAlaAsnThrMetGIyTrpSerPhe 349
1028 GCGGCGCTTCTCACCCTGCTGCCAGACCCAAACCGCGGCGGACCT 1077
349 rSerIlePheLeuThrValIleLeuProAspProLysProGlnGIyProPro 365
1078 GTGGCTCCGCTGTCGCGGCACTAGCTGCGCTGGCGCGGTGCTCATCG 1127
366 ValAlaSerSerSerSerAlaThrSerLeuProTrpProValIleGln 382
1128 CATCCAGCGCGGCTGTCTTCACTGGGACACCTGCTCTGTGGCTTT 1177
382 yIleProAlaGIyAlaValIlePheIleLeuGIyThrLeuLeuLeuTrpLeu 399
1178 GCCAGGCGCGAAGAACCGGTGACCCCGCGCTGCGCTCCCTGCGCT 1227
399 yseIleAlaGlnLysLysProCysThrProAlaProAlaProLeuPro 415
1228 GGGCAGCGCGCGCGGAGCGCGCGGACCGCGGAGCGAGACAGACCT 1277
416 GIyHisArgProProGIyThrAlaArgAspArgSerGIyAspLysAspLe 432
1278 TCCTGTGTGGCGCGCTCAGCGCTGCGCTGCTGTGGGCTGTGAGG 1327
|||||

```

```

432 uProSerLeuAlaAlaLeuSerAlaGlyProGIyValGIyLeu..... 446
1328 AGCATGGGTCTCCGGGACGCGCCGACACTTACTGGGCGGAGCGCCAGTT 1377
446 ..... 446
1378 GCTGGCCCTAAGTTGTACCCCAAACTCTACAGACATCCACACACAC 1427
446 ..... 446
1428 ACACGTGATTCGCGCGCGCTGTGTGAGAGACATGGTCTCCGACGCC 1477
447 .....Cys..GIyGlnHisGIySerProAlaIleArg 456
1478 CCGACCTTACTGGCGCGGCGGCGGCTGTGGCGCTGAGTTTACCCCA 1527
456 oGlnHisLeuLeuGIyProGIyProValAlaGIyProLysLeuTrpProL 473
1528 AACCTTACAGACATCCACACACACACACACACACATCTCACACACAC 1577
473 yseLeuTrpThrAspIleHisThrHisThrHisThrHisSerHisThrHis 489
1578 TCACAGTGGAGGCAAGGTCCACACAGACATCCACTATCATGTCG 1622
490 SerHisValGIyLysValIleGlnHisGlnHisIleHisTrpGIyCys 504

seq_name: sp_human:Q9NMWZ1

seq_documentation_block:
ID Q9NMWZ1 PRELIMINARY; PRT; 123 AA.
AC Q9NMWZ1;
DT 01-OCT-2000 (TrEMBLrel. 15, created)
DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, last annotation update)
DE CDNA FLJ20523 F1S, CLONE KAT10456.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Oobayashi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isono T., Sugano S.;
RT "NEO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AK000530; BAA91234.1; -.
SQ SEQUENCE 123 AA; 13282 MW; 62DA7D486EE4D1B CRC64;

alignment_scores:
Quality: 691.00 Length: 123
Ratio: 5.618 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-598-042a-2 x Q9NMWZ1 ..
Align seg 1/1 to: Q9NMWZ1 from: 1 to: 123

2659 ATGAGGAGCTTCCCTGCTCCACGCTGACCTCCGCAAGCTGCGCGCTC 2708
|||||
1 MetArgAspPheProCysSerThrValThrProProThrLeuProAlaAs 17
2709 TGTCCCGGCTCAGTCCGCGCTCATCCCGGCTGTGCCCTGCGCGCTT 2758
|||||
17 rValProAlaSerValProAlaSerIleProAlaSerValProTrpPro 34
2759 GCGGCTATTTTGGCAGCTGCTTGGGTGCCAGAGAGTCCCTACTGCT 2808
|||||
34 rPArgLeuPheLeuProProAlaLeuGIyAlaGlnIleSerProThrAla 50
2809 GTGGCTGGGGTTGGGGGCAAGCAGCAGCCCAAGCTTGAGAGGCTGAGGC 2858
|||||

```











```

151 uArgMetAspLysLysLeuLeuAlaValProAlaAlaAsnThrValArgp 168
491 TCAAGTGGTGGCCAGCGGCGACCCCTGGCCCGACATCATGCGTGAAG 540
168 heArgcysProAlaAlaGlyAsnProThrProSerTLeSerTrpLeuLys 184
541 GAGCAGCAGCGCCTTGACG.....CGCCAGAGCGCGCTGAGCCGAG 581
185 AsnGlyLysGluPheArgGlyGlnHisArgTLeGlyTLeLysLeuAr 201
582 GAAAGAGAAGTGAACACTGACCTGAAGAACCTGCGCGGAGAGACAGC 631
201 gHisGlnGlnTrpSerLeuValMetGluSerValValProSerAspArg 218
632 GCAATATACACTGCCCGCTGTGACACCGCGGCGCCATGACGCCAC 681
218 LysAsnTrpThrCysValValGlnLysAsnLysPheGlySerTLeArgGlnThr 234
682 TACAAGGTGATGATCGACGCGGACCGCTCCAAAGCCCGCTCACAG 731
235 TyrThrLeuAspValLeuGluArgSerProHisArgProTLeuGlnAl 251
732 CAGCAGCCCGCTGTAACAGCAGGTGACTTGGGGGAGACAGCTCTTC 781
251 AGlyLeuProAlaAsnGlnThrAlaTLeuGlySerAspValGlnPheH 268
782 AGTCCAGAGTGGCAGCGACAGCTGAAGCCGCTGATCCAGTGGCTGAAGCGC 831
268 LscYsLysValTyrSerAspAlaGlnProHisTLeGlnTrpLeuLysHis 284
832 GTGAGTATGCGCGCGAGCGCGGCACAACTCCACATCATGTCGGCGG 881
285 ValGlu.....ValAsnGln 289
882 CCAGAGATTGTGGTGGTGTCCAGCGGTGACGTGTGTGCGCGCGCAGC 931
289 ySerLysVal.....GlyProAspG 296
932 GCTCTCACTCAATTAAGCTGCTCATCACCCGTGCCCGCAG..... 972
296 LyrThrProTyrAspThrValLeuLysThrAlaGlyAlaAsnThrThrAsp 312
973 .....GACGATGCGCGG 983
313 LysGluLeuGlnValLeuSerLeuHisAsnValThrPheGlnAspAlaGln 329
984 CATGTACATCTGCTGGCGCCAAACACATGGGTACAGCTTCGCGACGG 1033
329 yGluTyrThrCysLeuAlaGlyAsnSerTLeGlyPheSerHisHisSerA 346
1034 CTTTCTCTACCGTGTGCGCA...GACCCAAAACCGCCAGGCGCACTGTG 1080
346 LArTrpLeuValValLeuProAlaGlnGlnGluLeuMetGluThrAspGlu 362
1081 GCCTCCTCGTCTGCGGCACACTAGCCTGCCGCGCTGTCATCGGCAT 1130
363 AlAGlySerValTyrAlaGlyValLeuSerTyrGlyValValPhe.... 377
1131 CCCAGCGCGCGCTGCTTCATCTGAGCGACCTGCTGCTGAGCTTGGC 1180
378 .....PheLeuPheTLeuValValAlaAlaValTLeuLysA 391
1181 AGGCCAGAGAGACGCTGACACCCCGCGCTGCCCTCCCTG..... 1224
391 rGluArgSerProProLysLysGlyLeuGlySerProThrValHisLys 407
1225 .....CCTGGGACCGCGCGCGCGGAGCGCCCGCGACGAG 1262
408 ValSerArgPheProLeuLysArgGlnValSerLeuGlnSerAspSe 424
1263 CGGAGACAGAGACCTTCCC.....TCGTTGACCGCGCTCAGCGGTGGCC 1306
424 rMetAsnSerAsnThrProLeuValArgTLeAlaArgLeuSerSerGlyG 441

```

```

1307 CTGCTGTGGGCTGTGCT.....GAGAGACATGGGTCTCCGCGACACCCC 1350
441 LngLyProValLeuAlaAsnValSerGluLeuGluLeuProAlaAspPro 457
1351 CAG 1353
458 Lys 458

seq_name: sp_vertebrate:Q91287
seq_documentation block:
ID Q91287; PRELIMINARY; PRT; 796 AA.
AC Q91287;
DT 01-NOV-1996 (TREMBLREL. 01, Created)
DT 01-NOV-1996 (TREMBLREL. 01, Last sequence update)
DT 01-MAR-2001 (TREMBLREL. 16, Last annotation update)
DE FIBROBLAST GROWTH FACTOR RECEPTOR 3.
OS Pleurodeles waltlilii (Iberian ribbed newt)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandridae; Salamandridae;
OC Pleurodeles.
OX NCBI_TaxID=8319;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94299014; PubMed=8026621;
RA Shi D.L., Launay C., Fromentoux V., Feige J.J., Boucaut J.C.;
RT "Expression of fibroblast growth factor receptor-2 splice variants is
RT developmentally and tissue-specifically regulated in the amphibian
RT embryo."
RL Dev. Biol. 164:173-182(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94274759; PubMed=8006062;
RA Shi D.L., Fromentoux V., Launay C., Umbhauer M., Boucaut J.C.;
RT "Isolation and developmental expression of the amphibian homolog of
RT the fibroblast growth factor receptor 3."
RL J. Cell Sci. 107:417-425(1994).
CC -1- SIMILARITY: TO IMMNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL: X75603; CAA53271.1; -.
DR HSP: P06239; 3LCK
DR InterPro: IPR000719; -.
DR InterPro: IPR001245; -.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 1g; 3.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR SMART: SM00219; TYRKC; 1.
KW ATP-binding; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 796 AA; 88288 MW; 226D99A0B6D1D2D CRC64;

alignment_scores:
Quality: 486.50 Length: 442
Ratio: 1.886 Gaps: 14
Percent Similarity: 58.371 Percent Identity: 31.222

alignment_block:
us-09-598-042a-2 x Q91287 ..
Align seg 1/1 to: Q91287 from: 1 to: 796

112 GCGCCATATAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 161
113 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
49 GlyAspThrTLeGlnLeuSerCysThrThrProGlySerSerValSerVal 65
162 GACCATGTGAGCAAGATGGCGCGACCATC.....C 193
65 L...ValTrpPheLysAspGlyTLeSerValAspProThrThrTrpSerH 81

```

```

194 ACAGCGCTGAGCGCGCTTCCGCGTGCAGGAGCGGTGAAGTGAAG 243
    |||||
81 1stIngly.....GlnlyLeuLeuLysIlelle 90
244 CAGGTGAGCGCGGAGAGATGCCGCGTGTACGTGCAAGGCCCAACG 293
    :|||:
91 AsnValSerTyrAspAspSerGlyValTyrSerCysLysAlaArgLys 107
294 CTTGCGGCGCTGAGCGTCACTACACCTGCTGTGTGTGATACATTA 343
    :|||:
107 rSerGluValLeuArg...AsnValThrValArgValThrAsp..... 120
344 GCCCAGGAGAGAGAGCGCTGGGCGCCGACACCTCTGTGGGGTCAAG 393
    :|||:
121 .....SerProSerSerGlyAspAspGluAspAsp 130
394 GACCCCGCCAGCCAGCACTGGGCGACGCGCGCTTCAACAGCCCTCAA 443
    |||||
131 AspGluGluSerGluSerAlaAsnAlaProLysPheThrArgProGluTr 147
444 GATGAGCGCGCGGTGATTCGACGCGCGGTGAGCTCGGTGGGCTCA 493
    |||||
147 pMetGlnLysLysLeuLeuAlaValProAlaAlaAsnThrValArgPhe 164
494 AGTCCGTGGCGCGCGCGCGCGCGCGCGCGCGCATCAGTGAAGAGAC 543
    :|||:
164 rGysProAlaAlaGlyLysProThrProSerIleThrPLeuLysAsn 180
544 GACCAGGCGCTTGAGC.....CGCCAGAGCGCGCTGAGCCAGGAA 584
    :|||:
181 GLyLysGluPheLysGlyGlnHisArgIleGlyLysLeuArgHis 197
585 GAAGAAGTGAACACTGAGCTGAAGAAGCTGCGCGGAGAGACGCGCA 634
    :|||:
197 sGlnGluTrpSerLeuValMetGluSerValProSerAspArgGly 214
635 AATGACACTGCGCGGTGTCAGACGCGCGCGCGCGCATCAAGCCACTAC 684
    :|||:
214 snYrTrpTrpCysValValAlaAsnLysTyrGlyThrIleArgGluThrTyr 230
685 AAGGTGATGTGATCCAGCGCGCGCGGTCCAGCGCGGTGTCAGAGCAC 734
    :|||:
231 ThrLeuAspValLeuGluArgThrProHisArgProIleLeuGlnAlaG 247
735 GCACCGCGTGAACAGACGAGTGTGCGGGGAGACAGCTCTTCAGT 784
    |||||
247 yPheArgSerAsnLysThrValValGlySerAspValGluPheHisC 264
785 GCAAGTGGCGCGAGCGTGAAGCGGTGATCCAGTGGCGTGAAGCGCGT 834
    |||||
264 ySylValTyrSerAspAlaGlnProHisIleGlnTrpLeuLysHisVal 280
835 GAGTACGCGCGCGAGGCGCGCACAACTCCACATCATGATGTCAGGCG 884
    |||||
281 Glu.....ValAsnLysE 285
885 GAAGTT.....GTGTCGTGCGCCAGCGGTG 910
    :|||:
285 rLysPheGlyProAspGlyAsnProTyrValThrValLeuLysThrAlaG 302
911 ACGTGTGGTGGCGCGCGCGTCTTACCTCAATAAGTGGCTCATACAC 960
    :|||:
302 LysVal.....AsnThrSerAspLysGluLeuLysIleGlnPheLeuArg 316
961 CGTGGCGCGAGGAGCATGCGGCGATCATCTGCTGGCGCGCAAC 1010
    :|||:
317 AsnValThrPheGluAspAlaGlyGluTyrThrCysLeuAlaGlyAsn 333
1011 CATGGCTCAGAGTTCGAGCGCGCTTCTCAACGCTGTCAGACGCCAA 1060
    :|||:
333 rLleGlyTyrSerHisHisSerAlaTrpLeuThrValLeuPro...ProA 349

```

```

1061 AACGCCAGGCGCACCTGTGTGCTCTGTCGTGGCACTAGCTTCCG 1110
    |||||
349 LagluProValProAspValAspThrSerValSer..... 360
1111 TGGCCCGTGTGCATCCGATCCAGCGCGCGTGTCTTCACTCGGCGAC 1160
    :|||:
361 .....IleLeuAlaAlaAlaGlyCysValAlaValIleLeuValVa 375
1161 CTGTCTCTCTGTGCTTGTGCGCGCGCGAGAAAGCCG..... 1197
    :|||:
375 ILeIleIlePheThrTyrLys...MetLysMetProSerLysLysThrK 391
1198 .....TCACCCCGCGCGCTGCGCGCGCTGCGCGCGCGCGCG 1239
    :|||:
391 etAsnThrAlaThrValHisLysValSerLysPheProLeuLysArgGln 407
1240 CCGGCGAGCGCGCGCGCGCGAGACAGACAGACCTTCC..... 1281
    :|||:
408 ValSerLeuGluSerAsnSerSerMetAsnSerAsnThrProLeuValAr 424
1282 ....TCGTTGGCGCGCGCTGAGCGCTGCTGTGTGGGCTGTGAGG 1327
    :|||:
424 gLleThrArgLeuSerSerSerAspGlyProMetLeuAlaAsnValSerG 441
1328 AGCATGGTCTCCGCGAGCGCGCGCGAG 1353
    |||||
441 lueGluLeuProAlaAspProLys 449

```

seq\_name: sp\_human:Q9H9U3

```

seq_documentation_block:
ID Q9H9U3 PRELIMINARY; PRT; 307 AA.
AC Q9H9U3:
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE CDNA FLJ12547 FIS, CLONE NT2RM4000634.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hirokawa S., Ishii S., Kawai Y.,
RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,
RA Masuno Y., Kanehori K.,
RT "NEDO human cDNA sequencing project."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK022609; BAB14128.1;
SQ SEQUENCE 307 AA; 32780 MW; 4CC18ACD39BD3AC3 CRC64;

```

alignment\_scores:

```

Quality: 485.50 Length: 430
Ratio: 2.357 Gaps: 32
Percent Similarity: 47.907 Percent Identity: 32.093

```

alignment\_block:

us-09-598-042a-2/rev x Q9H9U3 ..

Align seg 1/1 to: Q9H9U3 from: 1 to: 307

```

2574 CATATCTGACACTGTGTGTGTCATGTGTGTCGCGAAGTGTGCGCAGCA 2525
    |||||
2 HisValCysLeuCysValTrpVal..... 9
2524 GCATATCTGACATGTGTGCGTGTGTATCCGCGACATATCTGACG 2475
    |||||
10 HisLysValCysValProValCysVal.GlyAlaHisLeuCysVal 25
2474 TGTGTGATGTCCAGACAGCATATCTGTGACATGTGTGTGTCAGGCA 2425

```



RT "Acquisition of the human adeno-associated virus type-2 rep gene by  
RT human herpesvirus type-6.";  
RL Nature 351:78-80(1991).  
RN [16]  
RP SEQUENCE FROM N.A.  
RC STRAIN-U1102, VARIANT A:  
RX MEDLINE-91374590; PubMed-1654446;  
RA Martin M.E.D., Nicholas J., Thomson B.J., Newman C., Honess R.W.;  
RT "Identification of a transactivating function mapping to the putative  
RT immediate-early locus of human herpesvirus 6.";  
RL J. Virol. 65:5381-5390(1991).  
RN [17]  
RP SEQUENCE FROM N.A.  
RC STRAIN-U1102, VARIANT A:  
RX MEDLINE-92333249; PubMed-1321206;  
RA Efsthliou S., Lawrence G.L., Brown C.M., Barrell B.G.;  
RT "Identification of homologues to the human cytomegalovirus US22 gene  
RT family in human herpesvirus 6";  
RL J. Gen. Virol. 73:1661-1671(1992).  
RN [18]  
RP SEQUENCE FROM N.A.  
RC STRAIN-U1102, VARIANT A:  
RX MEDLINE-92148942; PubMed-1310766;  
RA Geng Y., Chandran B., Josephs S.F., Wood C.;  
RT "Identification and characterization of a human herpesvirus 6 gene 1  
RT segment that trans activates the human immunodeficiency virus type 1  
RT promoter.";  
RL J. Virol. 66:1564-1570(1992).  
RN [19]  
RP SEQUENCE FROM N.A.  
RC STRAIN-U1102, VARIANT A:  
RX MEDLINE-93091236; PubMed-1333836;  
RA Gompels U.A., Carss A.L., Sun N., Arrand J.R.;  
RT "Infectivity determinants encoded in a conserved gene block of human  
RT herpesvirus-6.";  
RL DNA Seq. 3:25-39(1992).  
RN [110]  
RP SEQUENCE FROM N.A.  
RC STRAIN-U1102, VARIANT A:  
RX MEDLINE-92260671; PubMed-1374813;  
RA Neipel F., Ellinger K., Fleckenstein B.;  
RT "Gene for the major antigenic structural protein (p100) of human  
RT herpesvirus 6.";  
RL J. Virol. 66:3918-3924(1992).  
RN [111]  
RP SEQUENCE FROM N.A.  
RC STRAIN-U1102, VARIANT A:  
RX MEDLINE-92333248; PubMed-1321205;  
RA Thomson B.J., Honess R.W.;  
RT "The right end of the unique region of the genome of human herpesvirus  
RT 6 U1102 contains a candidate immediate early gene enhancer and a  
RT homologue of the human cytomegalovirus US22 gene family.";  
RL J. Gen. Virol. 73:1649-1660(1992).  
RN [112]  
RP SEQUENCE FROM N.A.  
RC STRAIN-U1102, VARIANT A:  
RX MEDLINE-93187613; PubMed-8383182;  
RA Ellinger K., Neipel F., Foa-Tomasi L., Campadelli-Plume G.,  
RA Fleckenstein B.;  
RT "The glycoprotein B homologue of human herpesvirus 6.";  
RL J. Gen. Virol. 74:495-500(1993).  
RN [113]  
RP SEQUENCE FROM N.A.  
RC STRAIN-U1102, VARIANT A:  
RX MEDLINE-93224882; PubMed-8385692;  
RA Gompels U.A., Carrigan D.R., Carss A.L., Arno J.;  
RT "Two groups of human herpesvirus 6 identified by sequence analyses of  
RT laboratory strains and variants from Hodgkin's lymphoma and bone  
RT marrow transplant patients.";  
RL J. Gen. Virol. 74:613-622(1993).  
RN [114]  
RP SEQUENCE FROM N.A.  
RC STRAIN-U1102, VARIANT A:  
RX MEDLINE-93389439; PubMed-8397282;  
RA Liu D.X., Gompels U.A., Nicholas J., Lelliott C.;  
RT "Identification and expression of the human herpesvirus 6 glycoprotein  
RT H and interaction with an accessory 40K glycoprotein.";  
RL J. Gen. Virol. 74:1847-1857(1993).  
RN [115]  
RP SEQUENCE FROM N.A.  
RC STRAIN-U1102, VARIANT A:  
RX MEDLINE-94025558; PubMed-7692666;  
RA Liu D.X., Gompels U.A., Foa-Tomasi L., Campadelli-Plume G.;  
RT "Human herpesvirus-6 glycoprotein H and L homologs are components of  
RT the gp100 complex and the gH external domain is the target for  
RT neutralizing monoclonal antibodies.";  
RL Virology 197:12-22(1993).  
RN [116]  
RP SEQUENCE FROM N.A.  
RC STRAIN-U1102, VARIANT A:  
RX MEDLINE-93331710; PubMed-7687803;  
RA Pellet P.E., Sanchez-Martinez D., Dominguez G., Black J.B., Anton E.,  
RA Greenamoyer C., Dambaugh T.R.;  
RT "A strongly immunoreactive virion protein of human herpesvirus 6  
RT variant B strain 229. Identification and characterization of the gene  
RT and mapping of a variant-specific monoclonal antibody reactive  
RT epitope.";  
RL Virology 195:521-531(1993).  
RN [117]  
RP SEQUENCE FROM N.A.  
RC STRAIN-U1102, VARIANT A:  
RX MEDLINE-93323202; PubMed-7687301;  
RA Pfeiffer B., Bernehan Z.N., Neipel F., Chang C.K., Tiratnapong S.,  
RA Chandran B.;  
RT "Identification and mapping of the gene encoding the glycoprotein  
RT complex gp82-gp105 of human herpesvirus 6 and mapping of the  
RT neutralizing epitope recognized by monoclonal antibodies.";  
RL J. Virol. 67:4611-4620(1993).  
RN [118]  
RP SEQUENCE FROM N.A.  
RC STRAIN-U1102, VARIANT A:  
RX MEDLINE-95146989; PubMed-7844567;  
RA Gompels U.A., Macaulay H.A.;  
RT "Characterization of human telomeric repeat sequences from human  
RT herpesvirus 6 and relationship to replication.";  
RL J. Gen. Virol. 76:451-458(1995).  
RN [119]  
RP SEQUENCE FROM N.A.  
RC STRAIN-U1102, VARIANT A:  
RX MEDLINE-94047392; PubMed-8230490;  
RA Dewhurst S., Dollard S.C., Pellet P.E., Dambaugh T.R.;  
RT "Identification of a lytic-phase origin of DNA replication in human  
RT herpesvirus 6B strain 229.";  
RL J. Virol. 67:7680-7683(1993).  
RN [120]  
RP SEQUENCE FROM N.A.  
RC STRAIN-U1102, VARIANT A:  
RA Nicholas J.;  
RT Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
RN [121]  
RP SEQUENCE FROM N.A.  
RC STRAIN-U1102, VARIANT A:  
RX MEDLINE-94118404; PubMed-8289364;  
RA Nicholas J., Martin M.;  
RT "Nucleotide sequence analysis of a 38.5-kilobase-pair region of the  
RT genome of human herpesvirus 6 encoding human cytomegalovirus  
RT immediate-early gene homologs and transactivating functions.";  
RL J. Virol. 68:597-610(1994).  
RN [122]  
RP SEQUENCE FROM N.A.  
RC STRAIN-U1102, VARIANT A:  
RX MEDLINE-94202284; PubMed-8151768;  
RA Schiene U., Neipel F., Schreiner D., Fleckenstein B.;  
RT "Structure and transcription of an immediate-early region in the human  
RT herpesvirus 6 genome.";  
RL J. Virol. 68:2978-2985(1994).

RN [23]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-0102, VARIANT A:  
 RX MEDLINE-94181269; PubMed-8134119;  
 RA Thompson J., Choudhury S., Kashanchi F., Donliger J., Berneman Z.,  
 Frenkel N., Rosenthal L.J.

alignment\_scores:  
 Quality: 485.50 Length: 420  
 Ratio: 2.380 Gaps: 28  
 Percent Similarity: 48.571 Percent Identity: 33.095

alignment\_block:  
 US-09-598-042a-2/rev x Q69566 ..

Align seg 1/1 to: Q69566 from: 1 to: 413

```

2601 CACGTGATGCGGTGTCCTCC..... 2581
      |||||||.....
62 HisValSerValArgValSerValArgValSerValArgValSe 78
2580 .....AGGCAGCATATCTGCACCTGTGTGCA 2553
      ||| .....
78 rValCysAlaArgValCysAlaArgValCysValCysAlaArgValCysV 95
2552 TGTGTGTCGGAGAGTGTGTGTCGCCAGCAGCATATCTGCATGTGCGTGC 2503
      :|||:.....
95 aLcysAlaAlaArgValCysValCysAlaArg...ValCysValCysAlaArg 110
      ||||| .....
2502 GTGTGTGTCGGAGCAGCATATCTGCATGTGTGTCATCTGCACGACAT 2453
      ||||| .....
111 ValCysAlaAlaArgValCysValCysAlaCysVal...CysAlaC 126
      ||||| .....
2452 ATCTGTGCACATGTGTGTGTGTGTCCAGCAATATCTGCACGTGTGTGATGT 2403
      ||||| .....
126 yslLeuCysValCysAlaCysLeu...CysValCysAlaCysLeuCysVal 141
      ||||| .....
2402 GAGCAGACATATCTGTGTGTGTGTGTCCAGCAGCATATCTGCGTGTGT 2353
      ||||| .....
142 CysAlaCysLeuCysValCysAlaCysLeu.CysValCysAlaCysLeuC 158
      ||||| .....
2352 GTGTGTGTCGGAGCAGCATATCTGCATGTGTGTGTGTGTCGGA...CAG 2306
      ||||| .....
158 yslVal.....CysAlaCysLeuCysValCysAlaCysLeu 169
      ||||| .....
2305 CATGTGTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2256
      ||||| .....
170 CysValCysValCysVal.....CysLeuCysValCysVal..... 181
      ||||| .....
2255 AGCATATCTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2206
      ||||| .....
182 .....CysLeuCysValCysValCysLeuCysValCys..... 192
      ||||| .....
2205 ATATCTGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2156
      ||||| .....
193 ..ValCysLeuCysValCysValCysLeuLeuLeu..ValCysValCysLeuCy 208
      ||||| .....
2155 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2106
      ||||| .....
208 sValCysVal.....CysLeuCysVal..... 215
      ||||| .....
2105 ATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2056
      ||||| .....
216 .....CysVal.....CysLeuCysValCysValCys..... 224
      ||||| .....
2055 ATCTGCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2006
      ||||| .....
225 LeuCyValCysVal.....CysLeuCysValCysValCysValCysCy 238
      ||||| .....
2005 CAGCAGCAGCATATCTGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1956
      ||||| .....
238 s.....ValCysValCysValCysValCysValCys.....V 249
  
```

```

1955 TCCTGTATGTCTGTGCATGTCTGTGTGCGGTGTGTGTGTGTGTGTGTGTGTGT 1906
      ||| .....
249 aLcysValCysValCysValCysLeuLeuLeu.....Cys 262
      ||| .....
1905 TTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1856
      ||| .....
263 LeuCyValCysLeuCyValCysLeuCyValCysValCysValCysValCysV 279
      ||| .....
1855 TGTGCTTCAGGAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1806
      ||| .....
279 aLcysLeu.....CysValCysLeuLeuLeu..... 289
      ||| .....
1805 AGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1756
      ||||| .....
290 .....CysValCysValCysValCysValCysLeuLeuLeuLeuLeuSerLe 304
      ||||| .....
1755 CAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1706
      ||| .....
304 uCysMet.CysMetCysMetCysMetCysMet..... 316
      ||| .....
1705 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 317
      ||| .....
1655 CCCCCGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1606
      ||||| .....
328 .....LeuLeu.....MetC 331
      ||||| .....
1605 GCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1556
      ||| .....
331 ysmetCysmetCysmetCysmetCysmetCysmetCysmetCysmetCys 347
      ||| .....
1555 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1506
      ||| .....
348 llecysllecysmetCysmet..... 354
      ||| .....
1505 AACTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1456
      ||| .....
354 ..... 354
      ||| .....
1455 CCTCAGCAGGCGCGCATATACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1406
      ||| .....
355 .....CysmetCysmetCysmetCysmetCys 363
      ||| .....
seq_name: sp_r0dent:Q63237
seq_documentation_block:
ID 063237 PRELIMINARY; PRT: 446 AA.
AC 063237;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, last annotation update)
DE HEPARIN-BINDING FIBROBLAST GROWTH FACTOR RECEPTOR 2 (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
OX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PROSTATE;
RX MEDLINE=93326167; PubMed=8333865;
RA Yan G., McBride G., McKeehan W.L.;
RT "Exon skipping causes alteration of the COOH-terminus and deletion of
RT the phospholipase C gamma 1 interaction site in the FGF receptor 2
RT kinase in normal prostate epithelial cells.";
RL Biochem Biophys Res Commun. 194:512-518(1993)
CC -!- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL: L19105; AAA02628.1; -.
DR InterPro: IPR003006; -.
DR InterPro: IPR003598; -.
DR Pfam: PF00047; 1g. 3.
DR SMART: SM00408; IgC2; 1.
  
```





```

Ratio: 1.753      Gaps: 17
Percent Similarity: 59.041      Percent Identity: 30.719

Alignment block:
US-09-598-042A-2 x Q9JHX9 ..

Align seg 1/1 to: Q9JHX9 from: 1 to: 800

76  CCGGGTCGAGCATTTCTGGCAGGTGGCCGGCTG..... 111
   |||  ::|||  |||:::||||:::|
25  ProGUlInArGUaValG1ArGUArGAlaAlaGUaValProGUlProGUlPr 41
112  .....  GCGCGCACTATGCGGCTCG 130
   |||  |||  |||:::||||:|
41  oSeRgInGUlInGUlInAlaAlaPhnGUlSeRcUlsPrThrValGUlLeuS 58
131  AGTGC...CCAGTGGAGGGGAGCCCGCCGCGCTACATGTGGACCAAG 177
   ::|||  |||  |||  |||  |||:::||||:|
58  eRcYshSPRoProGUlG1ArGUaProthnGUlProthnLeuPrAlaYs 74
178  GATGGCCGCAACATCCACAGCGGTGAGCGCTTCCGGCTGCTGCCGA 227
   |||||  ::  |||  |||  |||  |||  |||
75  AspGUlYalGUlLeuValAlaSeR...HisArGUlLeuValGUlProGUl 90
228  GGGGCTAAGGTGAACAGAGGTGACCGGGAGATCCCGCGCTGACGT 277
   |  |||::|  ::::|  |||::|  |||::|  |
90  nArGUlGUlInValLeuGUaSnAlaThnHisGUlSpAlaGUlValYUSeRc 107
278  GC.....AAGGCCACCAAGCGCTTCGGCAGCGCTACAGGTCAACATACC 321
   ||  ::::  |||:::  ::::  ::::  ::::  ::::
107  YSGInGUlInArGUlThArGU.....ArGUaValLeuCUshSPReSeR 120
322  CTCGTCTGCTGTGATGCATTTAGCCCGAGGAAGAGAGCTGGGCCGA 371
   ::::  |||  |||  |||  |||  ::::  |||  ||
121  ValArGUaValThrArPaAlaProSeRSeRcUlsPaSPaGUlAspGUlYUs 137
372  CAGCTCTCTGGGGGTCAAGAGACCCCGCCAGCAGTGGGCGACAGAC 421
   |  ::  |||||  ::::  |||
137  PaValAla.....GUaSPThGUlYalProYUTrP..... 147
422  CGCGCTTCACACAGCCCTCCAAAGATGAGCGCGGGGTATGCCAGCGCC 471
   |||:::||||:::||||  :::::||||  |||  |||
148  .....ThArPProGUlInArGUeValSPYsYsLeuLeuAlaValPro 161
472  GTGGGTAGCTCCGTGGCGCTCAAGTGCCTGGCCAGCGGCGACCTCGGCC 521
   :::::||||:::||||:::||||  |||  |||  |||  ||
162  AlAlaInSnThrValArGUPhnArGUcYsPrAlaAlaGUlAsnProthPr 178
522  CGACATCAGCTGATGAAGAGCAGCACAGCGCTTGACG.....CGCG 562
   ||::||  |||:::||||:|  ::  ::  |||
178  oSeRlThProRPrLeuYUsAnGUlYUGlUPhArGUlGUlInHisArGUl 195
563  CAGAGCGCGCTGAGCCCGAGGAAGAAGTGAACCTAGACCTGAAGAAC 612
   ::::  ::  |||  :::::||||:|  :::::|
195  lEgUGlYlLeYsLeuArGUhInGUlInTPReSeRLeuValMetGUSeR 211
613  CTCGGCGCGGAGAGCAGCGCAATATACATCTCCCGCTGTCCAGCGCGC 662
   ::  |||::|  |||:::|||||  |||:::||||:|
212  ValValProSeRSPaRPaGUlAsnThrCUshValAlaGUlAsnYUsPh 228
713  CCAAGCCCGTCTACAGAGCAGCACCGCCGTGACACAGCAGCGTGAAC 762
   ::|||::|  ::  |||::|  |||::  |||::  ::
245  IsArGPProIlLeUGlInAlaGUlLeuProAlaInGUlInThrAlaValLeu 261
763  GGGGGGAGCAGCTCTCCAGTGCAGAGGTGGCGCAGCAGCTGAAGCGGT 812
   |||::|  :::::||||  |||  |||  |||  |||  |||
262  GUlSeRSPaRPaGUlInPhnHisCUshYsValYUGeRSPaRPaGUlInProh 278
813  GATGCCATGGCTGAAGCGCTGTGGAGTACGGGCGCAGAGGCCCGCCACACT 862

```

DR Pfam; PF00069; kinase; 1.  
 DR PRINTS; PRO0109; TYRKINASE.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS0011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR SMART; SM00219; TYRC; 1.  
 KW ATP-binding; Transferase; Tyrosine-protein kinase.  
 SQUENCE 802 AA; 89515 MW; CC5E5DF3BD25BD3 CRC64;

## alignment\_scores:

Quality: 475.00 Length: 445  
 Ratio: 1.813 Gaps: 17  
 Percent Similarity: 58.876 Percent Identity: 29.888

## alignment block:

US-09-598-042a-2 x 042127 ..

Align seg 1/1 to: 042127 from: 1 to: 802

```

109 CTGGGCGGCACTATGCGGCTGACATGCGGCAATGGAGGGGAGCCGCCGCC 158
    ::::::::::::::::::::::::::::|::|::|::|::|::|::|::|::|
44  lIeelygluThrlIeelylueSerCysAlaIaIu...AspAlaSerTh 59
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
159 GCTGACCATGTGACCAAGATGCGCGCCACATCCACAGCGCGTGAGCC 208
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
59  rThrlThrlSerCysAlaIaIu...AspAlaSerTh 67
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
209 GCTTCCCGCTGCTGCG...CAGGGG...CTG 234
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
68  ..lIeelylIeValAlaProAlaAlaArgThrlSerThArgGlnGlyLeu 83
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
235 AAGGTGAGCAGTGTGACGCGGAGATGCGGCTGACGCTGACGTCAGAGC 284
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
84  lYslIeIeAlaValSerSerAspSerCylIeIeIeSerCysAlaIe 100
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
285 CACCAACGCGCTTGGCAGCCTGACGCTCACTACACCTGCTGCTGCTG 334
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
100 uThrlIeSerThrlIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 116
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
335 ATGACATTAGCGCGGAGAGAGAGCTGGCGCGCGGACAGCTCTCTGCG 384
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
116 SP.....LeuProSerSerIeIeIeIeIeIeIeIeIeIeIeIeIeIe 127
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
385 GGTCAAGAGGACCCCGGCGGAGAGAGAGCTGAGGAGAGCGGCTGACACA 434
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
128 AspAspAspAspAspAspAspAspAspAspAspAspAspAspAspAsp 144
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
435 GCGCTCCAAATGAGCGCGCGGCTGATGCAAGCGCGCGGCTGATGCTG 484
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
144 nProGluArgMetIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 161
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
485 TGGGCTCAAGTGGTGGCGGAGAGAGAGCTGGCGCGGAGAGAGAGCTG 534
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
161 lEaIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 177
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
535 ATGAGGAGGAGGAGCGCTGAGC.....CGCCAGAGGCGCGCTGA 575
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
178 lEaIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 194
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
576 GCCCAGAGAGAGAGAGAGCTGAGCAGCTGAGAGAGAGAGAGAGAGAG 625
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
194 sLeuAlaIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 211
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
626 ACAGCGGCAATACACTGCGCGGCTGCGAAGCGCGGCGGCGGAGCTAAC 675
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
211 sPluIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 227
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
676 GCCACCTACAGAGTGTGATGATGACAGCGGAGCGGCTGACAGCGGCTG 725
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
228 GluThrlIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 244
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
726 CACAGGAGGAGGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 775
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

244 uGlnAlaIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 261
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
776 CTTTCAGATGCAAGTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 825
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
261 lUpHeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 277
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
826 AAGCGGCTGAGTACAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 875
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
278 lYslIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 282
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
876 GGGGCGGAGAGAGTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 925
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
282 lAsnGlySerIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 297
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
926 CCGAGCGGCTGACCTCAAT.....AAGCTGCTCATC 957
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
297 aLeuGlnSerPheIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 313
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
958 ACCGTCGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1007
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
314 lYslIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 330
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1008 CACCATGAGCTACAGCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1057
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
330 nPheIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 344
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1058 CAAGAGAG...CAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1104
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
345 ..lYslProAlaIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 360
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1105 CTGCGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1154
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
361 SerSerIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 377
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1155 GGGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1204
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
377 u...LeuValIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 393
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1205 CCGCG...CTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1233
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
393 ySerMetThrlAlaProIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 409
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1234 CGCGCGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1274
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
410 ArgGlnGlnValSerIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 426
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1275 CTTTCCCTGCTTGGCGGCGGCTGAGCGCTGCGCTGCTGCTGCTGCTG 1323
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
426 oLeuValArgIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 443
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1324 ....GAGGAGCATGGGTCTCCGCGGAGAGAGAGAGAGAGAGAGAGAG 1353
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
443 snValSerIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIeIe 454
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

seq\_name: sp\_vertibrate:Q91897

seq\_documentation\_block:

ID Q91897 PRELIMINARY; PRT; 814 AA.  
 AC Q91897;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE FIBROBLAST GROWTH FACTOR RECEPTOR PRECURSOR.  
 GN XIGFR OR E59.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Friesel R.E., David I.B.,

```

174 ySPrOAlaSmnIlythPrOserPrOAlaLeuAArgTlyleuLysmny 190
547 CAGGCGTTGACCGCCCAAG.....GCCGTGAGCCCAAGAA 584
    ::::: |||||::: ::::: |||:::
191 LysGIuPhe..ArgPrOAspIlnArgIleGIyGIyTlyLysValArgSe 206
    ::::: |||||::: ::::: |||||:::
585 GAAGAAGTGGACACTGAGCGCTGAAACACCTGCGCGCGGACAGCACGCGCA 634
    ::::: |||||::: ::::: |||||:::
206 rOlnIthrPrSerIleuIleMeLAsPserValValPrOserAspLysGIyA 223
    ::::: |||||::: ::::: |||||:::
635 AAPTACACCGCGCGGTGACGACCGCGGGCGGCATACAGCCCAAC 684
    ::::: |||||::: ::::: |||||:::
223 snrTyThrcysIleValIGuAsnLysTYrGIyThrLeuAsnIsthrTYr 239
    ::::: |||||::: ::::: |||||:::
685 AAGGAGATGATGATACAGCGAGACCGGTCCCAAGCCCGTGCTACAGGAC 734
    ::::: |||||::: ::::: |||||:::
240 GlnLeuAsPvalValGIlnArgSerPrOnIAsArgPrOIIleuGlnAlaGI 256
    ::::: |||||::: ::::: |||||:::
735 GCACCCCGTGAACAACAGCGGTGACTTCGGGGGACACAGTCTTCAGT 784
    ::::: |||||::: ::::: |||||:::
256 yLeuPrOAlaAsnIthrSerValIValGIySerIthrAlaGIuPheSerC 273
    ::::: |||||::: ::::: |||||:::
785 GCAAGGTGGCGACGCACGtGAAAGCGGGTATCCAGTGGCTGAAGCGCGT 834
    ::::: |||||::: ::::: |||||:::
273 yLysValTyLysrAsPPrOgIlnPrOnIstIleGIlnThrPheuArgIstIle 289
    ::::: |||||::: ::::: |||||:::
835 GAGTACGGGGCGGAGCGGCCCAACATCCACATGATGTGGCGGCA 884
    ::::: |||||::: ::::: |||||:::
290 GlnIle.....AsnIlySerArgValAlaIAserAspIlyPh 301
    ::::: |||||::: ::::: |||||:::
885 GAAGTTTGTG...GTGCTGCCACGAGGTGACGTGTGTGGCGGCCGACG 931
    ::::: |||||::: ::::: |||||:::
301 ePrOtyValIGlnIleIlyThrAlaGIyAl.....AsnThrSera 316
    ::::: |||||::: ::::: |||||:::
932 GCTCTACCTCAATAGCTGCTATACCCCGTGGCGCCGACAGCATGCG 981
    ::::: |||||::: ::::: |||||:::
316 sPlyAsPmetIGlnValLeuIlyLeuArgAsnValThrPheGIuAsPAla 332
    ::::: |||||::: ::::: |||||:::
982 GGCATGTATCATCTGGCTGGCGCCAAACACAGAGGCTACAGTTCGCGAG 1031
    ::::: |||||::: ::::: |||||:::
333 GIyGIlnTYrThrcysLeuAlaIAsnserIleGIyIleSerIlyAsnIst 349
    ::::: |||||::: ::::: |||||:::
1032 CGCCTTCTCACCCTGCTG.....CAAGACCCAAACACCGCCAGGCGCAC 1075
    ::::: |||||::: ::::: |||||:::
349 rAlaIthrPheIthrValLeuGIlnValGIuAsPAsPlyPro..... 362
    ::::: |||||::: ::::: |||||:::
1076 cMGTGAGCTCTMGTCCTCGGCCACTAGCTGCGCTGGCGCGTGGTCATC 1122
    ::::: |||||::: ::::: |||||:::
363 .....AlaLeuLeuAlaSerPrOleuGIlnLeuGIlnIleIle 375
    ::::: |||||::: ::::: |||||:::
1126 GGCATCCACCGCGGGCTGTCTTCATCTCGGGACACCTGTCTCTGGCT 1177
    ::::: |||||::: ::::: |||||:::
376 ..TYrCysThGIyAlaIAlaPheValSerAlaMetValThrIleI 391
    ::::: |||||::: ::::: |||||:::
1176 TTGCACAGGCCACAGAAAGACCTGTGACCCCGCGCGCT..... 1212
    ::::: |||||::: ::::: |||||:::
391 eIlePheIlysmetLysIthPrSerLysIlySerAsPheAsnserGIln 408
    ::::: |||||::: ::::: |||||:::
1213 .....GCCCTTCCCTMGCTGGGACCGCGCCGCGCGGG 1245
    ::::: |||||::: ::::: |||||:::
408 euAlaValAlaIstLysLeuAlaIstIlePrOleuArgArgGIlnValThr 424
    ::::: |||||::: ::::: |||||:::
1246 AGGCGCGCGGACCGACGAGCGGAACAAGAC.....CT 1277
    ::::: |||||::: ::::: |||||:::
425 ValSerGIyAlaSerAsnSerArgMetIstSerGIyAlaIleLeuValAr 441
    ::::: |||||::: ::::: |||||:::
1278 TCCTCTGTGTGGCGGCTCAGCGCTGCGGCTGTGTGGGCTGTGTGAG 1327
    ::::: |||||::: ::::: |||||:::
441 gPrOserArgLeuSerSetSerGIyThrPrOmetLeuSerGIyValSerG 458
    ::::: |||||::: ::::: |||||:::
1328 AGCATGGTCTCCGCGACGCCCCGACAG 1353
    ::::: |||||::: ::::: |||||:::

```

458 lutyrgluleuprogluaspproary 466

seq\_name: sp.vertebrate:Q9PS96

seq\_documentation\_block:  
 ID Q9PS96 PRELIMINARY; PRT; 810 AA.  
 AC Q9PS96;  
 DT 01-MAY-2000 (TRENBLREL. 13, Created)  
 DT 01-MAY-2000 (TRENBLREL. 13, last sequence update)  
 DT 01-MAR-2001 (TRENBLREL. 16, last annotation update)  
 DE FIBROBLAST GROWTH FACTOR RECEPTOR TYPE 1.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 NCBI\_TaxID:8355;  
 RX MEDLINE:96032697; PubMed-7559402;  
 RA Gillespie L.L., Chen G., Paterno G.D.;  
 RT "Cloning of a fibroblast growth factor receptor 1 splice variant from  
 RT Xenopus embryos that lacks a protein kinase C site important for the  
 RT regulation of receptor activity."  
 RL J. Biol. Chem. 270:22758-22763(1995).  
 DR HSP; P06239; 3LCK.  
 DR InterPro: IPR000719; -  
 DR InterPro: IPR001245; -  
 DR InterPro: IPR003006; -  
 DR Pfam: PF00047; Ig\_3.  
 DR Pfam: PF00069; pkinase; 1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR SMART: SM00219; Tyrc; 1.  
 DR ATP-binding; Transferase; Tyrosine-protein kinase.  
 SQ SEQUENCE 810 AA; 90294 MW; AC97B87BEBE1E72 CRC64;

## alignment\_scores:

Quality: 467.00 Length: 437  
 Ratio: 1.756 Gaps: 16  
 Percent Similarity: 60.870 Percent Identity: 30.435

## alignment\_block:

us-09-598-042a-2 x Q9PS96

Align seg 1/1 to: Q9PS96 from: 1 to: 810

103 GCCCGGCTGGGCGCACTATGCGCTGCACTGCCAGTGGAGGGGACCC 152  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 44 Alaglnprogllyaspargilletlrleuglncysargleuarglunspva 60  
 153 GCCCGGCTGACCATGTGGACCAAGATGGCCGACCATCCACAGCGGCT 202  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 60 lginserlleasn...trpallylsasnlglyalgnleu...serlur 75  
 203 GGAGCCGCTTCGCGTGTGCGCGAGGGGCTGAAGGAGCAAGGTTGAG 252  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 75 hrnsnagtrhrargilletlrleuglncysargleuarglunspva 91  
 253 CGGAGAGATCCGCGCTGTACGTGTGCAAGGACCAAGCGCTTCGCGAG 302  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 92 Proglunspasnlglyaltryalacs...valthrasnlgly..... 104  
 303 CCGAGCGTCACTACACCTC.....GTGCTGCTGGATGACATT 343  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 105 .ProserArgThrtyrThrValleupheserValasnValseraspalal 121  
 344 GCCGAGGAGAGAGCGCTGGGCGGACGACGCTCTCTGGGGTCAAGAG 393  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 121 euproserrlaglunspaspaspplunspaspaspasnserserser 137

394 GACCCGCGACCGACAGTGGCA.....CGACCGGCTTCACAGCC 437  
 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 138 GlulysalalaglunsnserlysproasnargProleuttrserhnspr 154  
 438 CTCGAGATGAGCGCGGCTGTATCGACGCGCGGTGGTAGTCCGTGC 487  
 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 154 OGluysmetglulyslyleuhsalavalProalalalysThrVal 171  
 488 GCGTCAAGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 537  
 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 171 yspheargysProalasnlglyThrProthrProthrleuaglyleu 187  
 538 AAGAGACGACGCGCTTCAGCGCGCGGCGGCGGCGGCGGCGGCGG 578  
 |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 188 Lysasnlglyarglalnleuglinsarglnarglileglylytrylsva 204  
 579 CAGGAGAGAGAGTGGACACTGAGCTGAAGAACTGGCGGCGGAGACA 628  
 |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 204 LargserlnThrtrpserleulemetaspservalvalProserasp 221  
 629 GCGGCAATGACCTGCGCGGTGTCGACCGCGGCGGCGGCGGCGGCGG 678  
 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 221 yselysnlytrncysllevalglunsnlystryglalalileasnhs 237  
 679 ACCTACAGGTGATGATCATCAGCGGACCGCTTCACAGCGCGGTCTAC 728  
 |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 238 ThrtyrclneunspvalvalaluaryserProhlsargprolleuagl 254  
 729 AGCGACGACCGCGGTGACAGAGGCTGCTGGGGGAGCACGCTCT 778  
 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 254 nalaglyleuproalasnhrservalthrvalglythrthralaglu 271  
 779 TCGAGGCAAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 828  
 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 271 heserlyslsvaltyrseraspproglprohlsileglntrpleuarg 287  
 829 CGCGTGGAGTACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 878  
 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 288 Hsileglule.....Asnlglyserargvalalaseras 299  
 879 CGGCCAGAGCTTGTG...GTGCTGCCAGCGGTGACGTGTGCGCGGC 925  
 |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 299 pgllypneprotyrvalglulileuulysthrvalaglyval.....Asn 314  
 926 CGGAGCGCTCTACCTCAATACCTGCTCATCACCGGCGGCGGCGGAG 975  
 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 314 hrserasplysaspmegluvalleuhsileuargsnvalthrpieglu 330  
 976 GATGCGGCGATGTACATCTGCTGGCGGCGGCGGCGGCGGCGGCTT 1025  
 |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 331 Aspalaglyaltrhrtyrlyleuhsalalasnserlleglylleserh 347  
 1026 CGGAGCGCTTCTCACCCTGCTG.....CGAGCCCAAAACCGGCGAG 1069  
 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 347 shisserrlatrpleuthrvalleuysvalgluasparnlyspro.... 362  
 1070 GCGCACCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1119  
 |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 363 .....AlaleuLeuAlaserProleuGlule 373  
 1120 GTCATGCGGATCCAGCGCGGCTGTTCATCTGCGGACCTGCTGCT 1169  
 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 374 llele...TyrCysThrlyalalaphelvalserAlametalvalah 389  
 1170 GTGGCTTGGCAGGCGGAGAGAGAGCGGTGCACCCCGCGGCT..... 1212  
 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 389 rlellellelphelymetylshisproserlyslsaseraspheasns 406  
 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 1213 .....GCGCTCCCTGCTGGGCGGCGGCGGCGGCGGCGGCGG 1239  
 406 erglneulavalahlslyslsleualalysertlepro..... 418  
 1240 CGGAGGAGCGCGCGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1289

```

419 .....LeuArgArgGlnValSerGlyAsp.....Se 427
1290 CGCCCTCAGCCCTGCTGCTGGGGCTGTGTGAGAGCATGGCTTC 1339
      :::: ||| :::: ||||| ||| :::: ::::
427 rSerSerSerMetasSerGlyValIleLeuValArgArgLeuSerSers 444
1340 CGGAGCGCCGCC 1350
      :::::||||
444 ergGlyThrPro 447

```

seq\_name: sp\_rodent:Q61563

seq\_documentation\_block:

```

ID_061563; PRELIMINARY; PRT; 782 AA.
AC_061563;
DT_01-NOV-1996 (TREMBLrel. 01, Created)
DT_01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT_01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE_FIBROBLAST GROWTH FACTOR RECEPTOR 3, ACID BOX-DELETED ISOFORM.
GN_FGFR3 OR FR3.
OS_Mus musculus (Mouse).
OC_Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC_Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX_NCBL_Taxid=10950;
RN_|||
RA_SEQUENCE FROM N.A.
RP_Shimizu A., Seo M.;
RL_Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN_|||
RP_SEQUENCE OF 224-346 FROM N.A.
RX_MEDLINE=94209351; PubMed=7512569;
RA_Chellatah A.T., McEwen D.G., Werner S., Xu J., Ornitz D.M.;
RT_Fibroblast growth factor receptor (FGFR) 3. Alternative splicing in
RT_immunoglobulin-like domain III creates a receptor highly specific for
RT_acidic FGF/FGF-1.
RL_J. Biol. Chem. 269:11620-11627(1994).
CC_-1 SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC_DOMAIN.
DR_EMBL: AF024638; AAB81604.1; -.
DR_EMBL: L26492; AAA21491.1; -.
DR_HSSP: P12931; IFMK.
DR_MGD: MGI:95524; FGFR3.
DR_InterPro: IPR000719; -.
DR_InterPro: IPR001245; -.
DR_InterPro: IPR003006; -.
DR_Pfam: PF00047; Ig; 3.
DR_Pfam: PF00069; pkinase; 1.
DR_Pfam: PR00109; TYRKINASE.
DR_PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR_PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR_PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR_SMART: SM00219; TYRKC; 1.
KM_ATP-binding; Tyrosinase; Tyrosine-protein kinase.
SQ_SEQUENCE 782 AA; 85871 MW; 90262E44F0CFE49B CRC64;

```

alignment\_scores:

```

Quality: 464.00 Length: 471
Ratio: 1.719 Gaps: 16
Percent Similarity: 57.325 Percent Identity: 30.786

```

alignment\_block:

us-09-598-042a-2 x Q61563 ..

Align seg 1/1 to: Q61563 from: 1 to: 782

```

9 GGGCGAGATGGATGCTCCGGCGCCGATGGCCGGGATGTCTTCA 58
||| |||||::: ||||| :::: ||| ::::
17 GAlaLarThrSerGlnProGlyProGlyProGlnGlnArg.....ValValAr 31
: : ||||| ||| :::: ||| ::::
59 TCCAGTATGTCGTCCGGGTGAGCATTTTCGTGGCAG..... 99
: : ||||| ||| :::: ||| ::::

```

```

31 gArgAlaIaGlnValProGlyProGlnProSer.GlnGlnGlnGlnVal 47
100 GTGGCGCCGGCTGGGCGGACATAGGGGCTGCAGTGC...CCAGTGAAGG 146
::: ||| |||:::|||||::: ||| |||
48 AlapheGlySerGlyAspThrValGlnLeuSerGlyShisProProGlyG 64
GGACCGCGCGCGCTGACCATGTGACGACAGAGATGACCGCACCATTCACA 196
| ||| |||:::|||||::: ||| ::::
64 yAlaProThrGlnProThrValTTPAlaLysAspGlyThrGlyLeuValA 81
| ||| |||:::|||||::: ||| ::::
197 GCGGCTGGAGCGGCTTCGCGTGTCTGCCGACAGGGGCTGAAGTGAAGCAG 246
::: |||::: ||| ||||| |||::: ||| :::
81 Laser...HisArgIleLeuValGlyProGlnArgLeuGlnValLeuasn 96
GTGAGACGGGAGAGATGCCGGCTGTAGCTGTGCAG.....GCCAC 287
:::|||||:::|||||::: ||| ::::
97 AlaSerHisGlnAspAlaGlyValTyrSerGlnHisArgLeuThrAr 113
CAGCGGCTTCGCGACCGCTGAGCGTCACTACACCCCTGCTGCTGATG 337
::: :::: :::::|||||::: |||
113 gArgValLeuGlyShisPheSerValArgValThr..... 124
338 ACATTAGCCCGAGGAGAGAGACCTGGGGCCGACAGCTCTGTGGGGGT 387
388 CAGAGAGACCCCGCCAGCCAGCTAGTGGCAGACCGGCTTCACACAGCC 437
125 .....GlyAlaProTyrTyrThrArgPr 132
438 CTCGAGATGAGGCGCGGCTGATGTCAGCGGCGGCTGGTACGTCGCGC 487
|:::|||||:::|||||::: |||:::|||||
132 GAluArgMetAspLysLysLeuLeuAlaValProAlaAlaSerThrValA 149
GAGCTCAAGTGGCGTGGCGCAGCGGACCCCTGCGCCGACATCAGTGTATG 537
|:::|||||:::|||||::: |||:::|||||
149 rghPheArgCysProAlaIaGlnAsnProThrProSerIleSerThrLeu 165
538 AAGGACGACGAGGCTTGACG.....CGCCGACAGGCGCGCTGAGCC 578
|:::|||||::: :::: |||
166 LysAsnGlyLysGluPheArgGlyGlnHisArgIleGlyGlyLeuValS 182
CAGGAAGAAGATGGACGACGCTGAGCTGAGAGACCTGCGCGGAGAGACA 628
|||:::|||||:::|||||::: |||:::|||||
182 uArgHisGlnGlnThrSerLeuValMetGlnSerValValProSerAspA 199
rGlyLysnThrThrCysValValGlnAsnLysPheGlySerIleArgGln 215
629 GCGGCAATATACACTGTCGCGGCTGTGAAACCGCGGCGGCGCATCAAGCC 678
|:::|||||:::|||||::: |||:::|||||
199 rGlyLysnThrThrCysValValGlnAsnLysPheGlySerIleArgGln 215
679 ACCTACAAAGTGGATGTGATCCAGCGGACCGGTTCCAAAGCCGCTGTAC 728
|:::|||||:::|||||::: |||:::|||||
216 ThrTyrThrLeuAspValLeuGlnArgSerProHisArgProIleLeuG 232
729 AGGACGACGACCGCGTGAACAGACGAGTGTGAGTGGGGGACAGCTGCT 778
::: |||:::|||||::: |||::: |||::: |||::: |||::: |||
232 nAlaGlyLeuProAlaAsnGlnThrAlaIleLeuGlySerAspValGluP 249
779 TCCAGTGCAGAGTGGCGCAGCGAGCTGAAGCGGTATGATCCAGTGGTGAAG 828
|:::|||||:::|||||::: |||:::|||||
249 nHisLysValValTyrSerAspAlaGlnProHisIleGlnThrLeuLys 265
829 CGGCTGAGATGACGCGCGGAGGCGCGGACCAACTCCACCATGATGTGG 878
:::|||||::: |||::: |||::: |||::: |||::: |||::: |||
266 HisValGlnVal.....AsnGlySerLysValGlyProAs 277
CGGCGAGAAGTTTGG...GTGCTGCCACGAGGTGACAGTGTGCGCGC 925
||| :::: |||::: |||::: |||::: |||::: |||::: |||::: |||
277 rGlyThrProTyrValThrValLeuLysThrAlaGly.....AlaAsnT 292
926 CGGAGGCTCTTACCTCAATAAGTGTATCAGCCGTCAGCCGCGCGAGAGC 975
::: |||::: |||::: |||::: |||::: |||::: |||::: |||
292 hThrAspLysGlnLeuGlnValLeuSerLeuHisAsnValThrPheGln 308

```

```

976 GATCGGGGATGTACATCTGCTTGGCCCAACACATGGGCTACACCTT 1025
||||| ||| |||||:|||||:|||||
309 AsplaglylutyrlthrcysleualaglyanserileglyPheSerH1 325
1026 CCGAGGCGCTCTCTCAACCGGTGCA...GACCCAAACCGCCAGGCC 1072
|||||:|||||:|||||:|||||:|||||
325 SHISerAlaTrpLeuValValLeuProAlaGluGluLeuMetLut 342
1073 CACCTGTGCTCTCTCTCTCGGCACCTAGCCTGGCGTGGCGGTGTC 1122
|||||:|||||:|||||:|||||:|||||
342 hrAspGluAlaGlySerValTyrlaGlyValLeuSerTyrglyValVal 358
1123 ATCGGCATCCACGCGCGCTGTCTTCATCCGTCGACCTGCTGTG 1172
::: |||||:|||||:|||||:|||||
359 Phe.....PhleuPheTrileuValValAlaAlaVal11 370
1173 GCTTGGCAGGCCCAAGAAGCGGTGACCCCGCGCGCTGCCCTCCC 1222
|||||:|||||:|||||:|||||:|||||
370 eLeuCyArGleuArGSerProProlysLysGlyLeuGlySerProThrv 387
1223 TG.....CCTGGCACCGCGCGCGCGCGGAGGAGCGCCGC 1254
::: |||||:|||||:|||||:|||||
387 aLHsLysValSerArgPheProLeuLysArgGlnValSerLeuGluSer 403
1255 GACCGCAGCGGAGACAAGACCTTCCC.....TCGTTGGCGCGCTCAG 1298
::: |||:|||||:|||||:|||||:|||||
404 AsnSerSerMetAsnSerAsnThProLeuValArgLeuAlaArgLeuSe 420
1299 CGCGGCGCTGTGTGTGGCTGT.....GAGGACATGGGCTCTCGG 1342
||||| ||| |||||:|||||:|||||
420 rSerGlyGluGlyProValLeuAlaAsnValSerGluLeuGluLeuProA 437
1343 CAGCGCCCGCAG 1353
|| |||||
437 laAspProlys 440

seq_name: sp_human:Q14309

seq_documentation_block:
ID 014309 PRELIMINARY: PRT: 802 AA.
AC 014309: 043785.
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DE 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE FIBROBLAST GROWTH FACTOR RECEPTOR 4 (FGFR-4) (EC 2.7.1.112).
GN FGFR4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN RP TISSUE=MAMMARY GLAND;
RC TISSUE=MAMMARY GLAND;
RA Patanen J.M.;
RL Submitted (Oct-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RN RP TISSUE=MAMMARY GLAND;
RC TISSUE=MAMMARY GLAND;
RA Ron D., Tronick S.R., Lengel C.R.;
RL Submitted (Oct-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RN RP SEQUENCE FROM N.A.
RA Kozlowska M., Mueller U.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL: L03840; AAB59389.1; -.
DR EMBL: Y13901; CA474200.1; -.
DR HSPF: P06239; 3LCK.
DR InterPro: IPR000719; -.
DR InterPro: IPR001245; -.
DR InterPro: IPR003006; -.

```

```

DR Pfam: PF00047; 1g; 3.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR SMART: SM00219; Tyrc; 1.
KW ATP-binding; Glycoprotein; Receptor; Transferase;
KW Tyrosine-protein kinase.
FT CONFLICT 10 10 V -> I (IN REF. 3).
FT CONFLICT 136 136 P -> L (IN REF. 3).
SQ SEQUENCE 802 AA; 87954 MW; B2B259831BB889F CRC64;

alignment_scores:
Quality: 459.50 Length: 405
Ratio: 1.923 Gaps: 15
Percent Similarity: 59.012 Percent Identity: 31.852

alignment_block:
US-09-598-042A-2 x Q14309 ..
Align seg 1/1 to: Q14309 from: 1 to: 802

109 CTGGCGCGCACTATGCGGCTGTGACGCGCCAGTGAAGGAGCGCGCCGCC 158
|||||: ||| |||||: |||
49 LeuGlyLnpProValAlaGluCysCys.....GlyArgAlaGluAr 62
159 GCTGACCATGTGACCAAGATGCGCGCGCACCATCCACAGCGGTGAGACC 208
||| |||||: |||
62 gLyGlyLHsTrpTyrlGlyGluGlySer..... 71
209 GCTTCGCGCTGTCCCGCAGGCG.....CTG 224
|||||: ||| |||||: |||
72 ....ArgLeuAlaProAlaGlyArgValArgLtyrPArgLtyrGlu 86
235 AAGGTGAAGAGGTGAGCGGAGATGCCGCGGTGACGTGTGCAAGCG 284
|||||: ||| |||||: |||
87 GluLeAlaSerPheLeuProGluAspAlaGlyArgTyrlLeuGlySer 103
285 CACCAACGCTTGGCAGCTTGAAGCTC.....AACTACACCTGTGTCG 328
||| |||||: |||
103 aArg.....GlySerMetLeuValLeuGlnAsnLeuThrLeuLleT 117
329 TGCTGATGATCATTTACCCAGGAGAGAGAGCTGGCGCGCCAGCTGCG 378
|||||: ||| |||||: |||
117 hrcLysPsrSerLeuThrSerSerAsnAspAspLysProlysSer... 132
379 TCTGGGGGTCAAGAGACCCCGCC.....AGCCAGCATGGGCGACG 419
|||||: ||| |||||: |||
133 .....HisArgAspProSerAsnArgHisSerTyrlProGlnGln 146
420 ACCGCGCTTCACACACGCTCCCAAGTAGAGCGCGCGGTATGCGAGCG 469
|||||: ||| |||||: |||
146 aProTyTrpThrHisProGlnArgMetGluLysLysLeuHisAlaValAr 163
470 CCGTGGTGAAGCTCCGTGCGGCTCAAGTGGTGGCGCGCGCGCTGCG 519
|||||: ||| |||||: |||
163 roAlaLysAsnThrValLysPheArgCysProAlaAlaLysAsnProThr 179
520 CCGGACATCAGTGAATGAAGAGCAGCAGCGCTTG.....ACGCG 560
|||||: ||| |||||: |||
180 ProThrLArgTrpLeuLysAspGlyGlnAlaPheHisGlyGluAsnArg 196
561 CCGAGAGCGCGCTGAGCCGAGAGAGAGAGAGTGGACAGTGGAGCTGA 610
|||||: ||| |||||: |||
196 gLleGlyGlyLtyrLeuArgHisGlnHisTrpSerLeuValMetGluS 213
611 ACCTGGCGCGCGGAGCAGCGCAATACAGCTGGCGCGCTGCGAGAGCG 660
|||||: ||| |||||: |||
213 eValValProSerAspArgGlyThrTyrlThrcysLeuValGluAsnAla 229
661 GCGGCGCGCATCAAGCCACCTAAGAGTGAATGTGATCCAGCGGAGCGG 710

```

```

230 ValGlySerIleArgTyrAsnTyrIleuLeuAspValIleuGluArgSerPr 246
246 ValGlySerIleArgTyrAsnTyrIleuLeuAspValIleuGluArgSerPr 246
711 TTCACAGCCCTGCTCAGACGACGACCCGCTGACAGACGAGCTGACT 760
246 OHISArgProIleuGlnIleuGluArgProIleuAsnThrIleuVal 263
761 TCGGGGGGACACGCTCTTCCAGTGCAGAGTGCAGACGACGACGAGCG 810
263 alGlySerAspValIleuLeuGlyValTyrSerAspAlaGlnPro 279
811 GTGATTCAGTGCAGAGCTGAGTGCAGAGTGCAGAGTGCAGAGTGCAG 860
280 HistIleGlnThrIleuValIleuValIleuValIleuValIleuVal 291
861 CTCACACATCATGTGAGGCGCCAGAAATTTGTG...GTGCTGCCACAG 907
291 ySerSerIleuGluAlaAspGlyPheProTyrValGlnValIleuGlyThr 308
908 GTGACGTGTGCTGCGGCGCCGACGCTCTTACCTCATTAAGCTGCTCATC 957
308 IAspIle...AsnSerSerGluValGluValIleuGlyThr 320
958 ACCGTCGCGCCGACGAGGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 1007
321 ArgAsnValSerAlaGluAspAlaGluIleuGlyThrCysLeuAlaGlyAs 337
1008 CACCATGGGCTACAGCTCCGACGACGCTCTCTGACCTGCTGACGACG 1057
337 nSerIleuGlyLeuSerTyrGlnSerAlaThrIleuThrValIleuProGln 354
1058 CAAACCGCCGACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1107
354 IAspProThrTyrThrAlaAlaAlaProGluAlaArgTyrThrAsp... 369
1108 CCGTGGCGCCGCTGATGCGATCCGACGCGCTGCTGCTGCTGCTGCTG 1157
370 .....IleIleuThrAlaSerIleuAlaAlaAlaAlaAlaAlaAla 382
1158 CACCTGCTCTCTG...TGGCTTTCGACGACGACGACGACGACGACG 1204
382 IleuIleuLeuLeuAlaGlyLeuTyrArgGlyGln..... 393
1205 CCGGCGCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1248
394 .....AlaLeuHisGlyArgHisProArgProProAlaThr 405
1249 GCCCGCGACGACGACG 1263
406 ValGlnIleuSer 410
seq_name: sp_rodent:063236
seq_documentation_block:
ID 063236 PRELIMINARY; PRT; 446 AA.
AC 063236;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE HEPARIN-BINDING FIBROBLAST GROWTH FACTOR RECEPTOR 2 (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP .SEQUENCE FROM N.A.
RC TISSUE=PROSTATE;
RA MEDLINE=93326167; PubMed=8333865;
RA Yan G., McBride G., McKeehan W.L.;
RT "Exon skipping causes alteration of the COOH-terminus and deletion of
RT the phospholipase C gamma 1 interaction site in the FGF receptor 2
RT kinase in normal prostate epithelial cells.";
RL Biochem. Biophys. Res. Commun. 194:512-518(1993).

```

```

CC -I- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL, L19104; AAA02629.1; .
DR InterPro: IPR003006; .
DR InterPro: IPR003598; .
DR Pfam: PF00047; Ig; 3.
DR SMART, SM00408; IGG2; 1.
KW Alternative splicing.
FT NON_TER 1
FT NON_TER 446
FT NON_TER 446
SQ SEQUENCE 446 AA; 49379 MW; 276DF304DD3080E3 CRC64;

alignment_scores:
Quality: 458.50 Length: 378
Ratio: 1.959 Gaps: 9
Percent Similarity: 61.905 Percent Identity: 31.481

alignment_block:
US-09-598-042a-2 x Q63236 ..

Align seg 1/1 to: Q63236 from: 1 to: 446

112 GCGCGCACTATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 161
82 GlyGlySerIleuGlnIleuArgCysMetLeuLys.....AspAlaAlaVal 96
162 GACCATGTGACCAAGATGGCGCGACCATGACGACGCTGCTGACGCTGCT 211
96 IleSerThrIleuLysAspGly.....ValHisIleuGlyProAsnAsn 111
212 TCCGCTGCTG...CCGACGAGGCTGAAGGTGAAGTGAAGTGAAGTGAAG 258
111 rGthValIleuIleuGlyIleuGlyIleuGlyIleuGlyIleuGlyIleu 127
259 GATGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 308
128 AspSerGlyLeuTyrAlaCysAlaAlaAlaAlaAlaAlaAlaAlaAla 144
309 GGTACATACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358
144 rLeuTyrThrMetValAsnValThrAspAlaIleSerSerGlyAspAsp 161
359 GCTTGGGCGCGACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 408
161 IAspAspThrAspSerSerGluAspPheValSerGluAsnArgSerAsn 177
409 CAGTGGGACACGACGCGCTTCCACACGCGCTCCAGATGAGGCGCGGCT 458
178 GlnArgAla...ProTyrThrIleuAsnThrGlyIleuGlyIleuGly 193
459 GATGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 508
193 HisAlaValAlaProAlaAlaAsnThrValIleuPheArgGlyProAlaGly 210
509 GCGACGCTGCGCGCGACATCAGTGCATGATGAAGGACGACGACGCTTACG 558
210 LysAspProThrProThrMetArgTyrLeuLysAsnGlyLysIleuPheLys 226
559 .....CGCCGACGAGCGCGCTGACGCGCGAGAGAGAGAGAGAGAGACT 599
227 GlnGlnHisArgIleuGlyIleuTyrValIleuValIleuValIleuVal 243
600 GAGCTGAAGACCTGCGCGCGAGGACGACGCGCAATATACACTGCGCGCG 649
243 IleMetGluSerValIleuProSerAspLysIleuGlyIleuGlyIleuGly 260
650 TGTTCACCGCGCGCGCGCATCAAGCGACGACGCTGACGCTGCTGCTGCT 699
260 alGlnAsnGlyIleuGlySerIleuAsnHisThrTyrHisIleuAspValVal 276
700 CAGCGGACCGCTTCCAGCGCGCTGCTGACGAGGACGACGCGCGTGAACAC 749
:::|||||:::
:::|||||:::

```

```

277  GLuArgSerProHisArgProIleLeuGlnAlaGlyLeuProAlaAsnAl 293
750  GACGGTGGACTTCGGGGGACACAGCTCTTCAGTGCAGAGTGGCGACG 799
293  aserThrValValGlyGlyAspValGlnPheValCysGlyValTyrSer 310
800  ACGTGAAGCCGGTGAATCCAGTGGCTGAAGCCGCTGAG..... 837
310  spAlaGlnProHisIleGlnTyrPheValHisValGlnGlyAsnGlySer 326
838  ...TAGGGCCGGGAGGGCCGACAACTCCACCATCATGTGGGGCCCA 884
327  LysTyrGlyProAspGly..... 332
885  GAAGTTTGTTGCTGCTGCCACGGGTGACGTG.....TGGTCCGGCGCCG 928
333  .....LeuProTyrLeuLysValLeuLysHisSerGlyLeu 345
929  ACGGCTCTACCTCAATTAAGTGTCTATCACCCGTGCCCGCCAGAGCAT 978
345  snSerSerAsnAlaGlnValLeuAlaLeuPheAsnValThrGlnMetAsp 361
979  GCGGGCATGTACATCTGCTTGGGGCCACACCATGGGCTACAGCTCCG 1028
362  AlaGlyAluTyrIleCysLysValSerAsnTyrIleGlyGlnAlaAsnG 378
1029  CAGCGCTTCCTCAACGCTGCTGCCAGACCCAAACCCGAGGGCCACCTG 1078
378  nSerAlaTyrLeuThrValLeuProLysGlnGlnAlaPro..... 391
1079  TGGCTCTCTGCTGCTCGGCGCATAGACCTGCGCGCGGTGTCATCGGC 1128
392  ..ValArgGlnLysGlnIleThrAlaSerProAspTyrLeuGlnIleAla 407
1129  ATCCACCGCGGGCGGTGCTCTATCATCTGGGACACCTGCTGCTGCGTTG 1178
408  IleTyrGlnIleGlyValPheLeuIleAlaCysMetValValThrValI 424
1179  CCAAGGCCAGAGAAGACCGCTGCACCCCGCGCGCT 1212
424  ePheCysArgMetLysThrThrThrLysLysPro 435

seq_name: sp.vertebrate:Q9DGK3

seq_documentation_block:
ID Q9DGK3 PRELIMINARY; PRT; 828 AA.
AC Q9DGK3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE FIBROBLAST GROWTH FACTOR RECEPTOR 4C.
GN FGFR-4C.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8335;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed-11180841;
RA Golub R., Adelman Z., Clementi J., Weiss R., Bonasera J.,
RA Servetnick M.;
RT "Evolutionarily conserved and divergent expression of members of the
RT FGF receptor family among vertebrate embryos, as revealed by FGF
RT expression patterns in Xenopus.";
RL Dev. Genes Evol. 210:345-357(2000).
DR EMBL; AF288453; AAG01013.1; -.
KW Receptor.
SQ SEQUENCE 828 AA; 93211 MW; 72E2052635DE5323 CRC64;

```

alignment\_scores: Quality: 457.50 Length: 375

Ratio: 1.930 Gaps: 11  
Percent Similarity: 63.200 Percent Identity: 30.400  
alignment\_block:  
US-09-598-042a-2 x Q9DGK3 ..  
Align seg 1/1 to: Q9DGK3 from: 1 to: 828

```

112  GCGCGCATGTGGGGCTGCAGTCCAGTGCAGAGGAGCCGCCGCGCT 161
60  GLysAlaLeuArgLeuPheCys.....AspThrAsnGlnSerSer 74
162  GACCATGTGACCAAGAGATGCGGCACACATCCACAGCGCTGAGCGCGT 211
74  rIleAsnTyrTyrArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnG 212
212  TCCGCGTGGTCCCGCAGAGGCTGAAGTGAAGAGAGTGAAGAGAGAT 261
90  leArgMetValGlyThrValLeuGlnValSerAspValThrTyrGlnAsp 106
262  GCGCGCATGTGGGGCTGCAGTCCAGTGCAGAGGAGCCGCCGCGCTGAG 311
107  SerGlyLeuTyrIleCys..ValValArgGlyThrGlyLysIleLeu 122
312  CAACATACCCCTGCTGCTGCTGATGATGATGATGATGATGATGATGAT 352
122  gATgPheSerIleSerValValAlaSerLeuAlaSerGlyAlaGlnGln 139
353  AGAGAGACCTGGGGCCCGACAGCTCTGCGGGGTGAAGAGAGAGAGAG 402
139  IuGlnGlnGlnAlaAspAspAspGlnAspGlnAspGlnAspGlnAsp 155
403  AGCGAG.....CAGTGGGACAGACCGCGCTTAC 431
156  AlaAspIleAsnGlnIleLysProValTyrPhePheGlnAlaProTyr 172
432  ACAAGCCCTCCAGATAGAGCGCGCGGTGATGCGACGCGCGCTGGTAG 481
172  rGlnProHisArgMetAspLysLysLeuHisAlaValProAlaGlyAsn 189
482  CGGTGGGGCTCAAGTGCCTGGCGAGCGGCGAGCGCGCGCGCGAGCAG 531
189  hValLysPheArgCysProAlaGlyLysPheProLysProThrIleArg 205
532  TGGATGAAGAGCAGCAGCGCTTGAG.....CGCCAGAGGCGCG 572
206  ThrLeuLysAsnGlnLysArgGlnPheArgGlnHisArgIleGlyLys 222
573  TGAAGCCAGAGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 622
222  eGlnIleArgHisGlnHisTyrSerLeuValMetGlnSerValValPro 239
623  AGGACAGCGCCAAATACCTGCGCGGTGATGAGCGCGCGCGCGCGAGC 672
239  eAspArgGlyAsnTyrThrCysValValGlnAsnArgValGlySerLeu 255
673  AACGCCATACAGTGTGATGATGATGATGATGATGATGATGATGATGAT 722
256  ThrTyrThrTyrPheLeuAspValLeuGlnArgSerSerHisArgPro 272
723  GCTCAGAGCAGCAGCAGCGCTGATGATGATGATGATGATGATGATGAT 772
272  eLeuGlnAlaGlyLeuProAlaAsnThrThrAlaArgValGlySerAsp 289
773  CGAGCTTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 822
289  aGlnPheTyrCysLysValTyrSerAspAlaGlnProHisIleGlnTyr 305
823  CTGAAGCGCGTGAAGTGAAGCGCGCGAGCGCGCGAGCGCGAGCGAG 872
306  LeuLysHisIleGln.....ValAsnGlySerHis..... 315
873  TGTGGCGGCGCAGAGAGTTT.....GTGGTGTGCGCCAGCGGTGACG

```





```

347 Thrmctluaspalaiglytutyrthrcysleualaglynsertilegl 363
1017 CTACAGCTTCGCGAGCGCTTCCTCCACACCGCTGCTCCAGACCCAAACCGC 1066
363 ylleuSerHisglInsrtalatrpleuthrvalleuproasnglu..... 377
1067 CAGGGCCACCTGTGGCTCCTCTGCTGCGCCACTAGCTGCGGTGGCC 1116
378 ....AspHeuLeugluInlaagluProalagInsrtarArgTyrMetAsp 392
1117 GTGGTCACTGGCATCCACCGCGGCTGTCTTCATCTGGGACACCTGCT 1166
393 llellelletytrHsrerGlypheleualalavalametalallevalil 409
1167 CCGTGGCTTGGCAGGCGCCAG 1188
409 eValilleuLeucysArGmetcIn 416

```

seq\_name: sp-vertebrate:Q9PSV9

seq\_documentation\_block:

```

ID Q9PSV9 PRELIMINARY: PRT: 818 AA.
AC Q9PSV9
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE FGF RECEPTOR 4A.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OX Xenopodinae; Xenopus.
RX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Hongo I., Kengaku M., Okamoto H.;
RT "Differential employment of FGF signalling system for the embryonic
RT induction."
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB007036; BAA22849.1; -
DR HSSP; P06239; 3LCK.
DR InterPro; IPR000719; -
DR InterPro; IPR001245; -
DR InterPro; IPR003006; -
DR Pfam; PF00047; 1g; 3.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR SMART; SM00219; TyrcK; 1.
KW ATP-binding; Receptor; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 818 AA: 92067 MW; A8CBAA341C9230C1 CRC64;

```

alignment\_scores:

```

Quality: 455.50 Length: 396
Ratio: 1.922 Gaps: 10
Percent Similarity: 59.848 Percent Identity: 29.293

```

alignment\_block:

US-09-598-042a-2 x Q9PSV9 ..

Align seg 1/1 to: Q9PSV9 from: 1 to: 818

```

112 GCGCGCACTGTGGGCTGCGAGTGCACAGTGAGGGAGCCGCCGCCGCT 161
|||||
56 GllYasnlaaleuArGlyeupheCysAsp..... 64
162 GACCATGTGACCAAGATGGCGCACCATGCAACAGCGGCTGAGCGCGCT 211
|||||
65 .....ThrasnGlnserAsnserIleAsntrIPtyrArggluGln 78

```

```

212 TCCGCTGCTGCCGACGGG.....CTGAG 237
|||||
78 spArGlyleuProGllYglYlIeArGmetValGlyThrValleuGlu 94
238 GTAGACAGGTGGAGCGGAGAGATGCCGGCTGTAGCTGTGCAGGCCAC 287
|||||
95 ValSerAspValthrYtrGluAspSerGlyLeuTyrIleCys...ValVa 110
288 CAAGGCTTGGCGAGCGTGAAGCTCACTACACCTCCGTCGTGGATG 337
|||||
110 lArGgLYThrGlyLysIleleuArGArGpHeSerIleSerValValAspS 127
338 ACATTAAGCCAGGAAGAGAGAGCGTGGCGCCACAGCTCTCTGGGCT 387
|||||
127 erleuAlaserGlyAspGlu.....AspAspGluAspGlyArg 140
388 CAAGAGAGCCCGCCAGCCAG.....CAGTGGC 416
|||||
141 ArgGluAspThrThrAlaAspIleAsnGluGluProValTyrPhePhegl 157
|||||
157 nAlaProTyrTrpThrGlnProHisArGmetAspLysLysLeuHisAlav 174
467 GCGCGGTGGGTAGCTCCGTGCGCTCAAGTGCCTGGCCAGCGGACCCCT 516
|||||
174 alProAlaGlyAsnThrVallyPheArGcysProAlaGlyGlySerPro 190
517 CGCGCCGACATCACTGATGAAGAGACGACGACCTTGACG..... 558
|||||
191 LeuProThrThrleArGlyrPheLysAsnGlyArgGluPheArgGlyGluH 207
559 .CGCCAGAGCGCGCTGAGCCCGCAAGAAAGATGAGACACTAGCTGCA 607
|||||
207 sarGllleGlyGlyIleGlnleuArGHisGlnHsrPserLeuValMetc 224
608 AGAAGCTGGCGCGGAGAGACGCGCAATACACTGCGCGCTGTGCAAC 657
|||||
224 lAsrValValProSerAspArgIlyAsnTyrThrCysValValGluAsn 240
658 CGCGCGGCGCCCTCAACGCGCACTCAAGGTGATGTATCCAGCGGAC 707
|||||
241 ArgValGlySerLeuThrTyrThrYrPheLeuAspValleuGluArgSe 257
708 CCGTTCACAGCCGCTGCTACAGCAGCAGCCCGGAGAACAGCGGTG 757
|||||
257 rSerHisArGProIleleuGlnAlaGlyLeuProAlaAsnThrThrAla 274
758 ACTTGGGGGAGACAGCTCTTCACAGTGCAGAGGTGGCGACGAGTGAAG 807
|||||
274 rGValGlySerAspValGluPheTyrCysIlyValTyrSerAspAlaGln 290
808 CCGGTATCCAGTGGCTGAAGCGCGTGGAGTACGGCGCCGAGGCCGCCA 857
|||||
291 ProHisIleGlnThrPheLysHisIleGlu..... 300
858 CAAGCTCACCATGATGTGGCGCGCAGAGTT..... 891
|||||
301 .....ValAsnGlySerArGpHeGlyProAspAspPhe 312
892 .....GTGTGCTGCCAGCGGTGACGTGTGTGGCGCCGACGGC 933
|||||
312 rGlyrValGlnValleuLysThrAlaAspIle.....AsnThr 324
934 TCCTACTCAATAAGCTGCTCATCACCGGTGGCGCCAGAGCATGTGGG 983
|||||
325 SerGluValGlnValleuHisLeuArGAsnIleThrMetcLusAspAlaG 341
984 CATGTACATGCTGCGCGCGCAACACATGAGGCTCAAGTTCGGAGCG 1033
|||||
341 yGluTyrThrCysLeuAlaGlyAsnSerIleGlyLeuSerHisGlnSerA 358
1034 CCTTCTCACCGGTGCTGCCAGACCCAAACCGCGAGGCGCACCTGTGGCC 1083

```

```

1133 358 1atrpLeuThrValLeuSerAsnGlu.....AspPheLeuGlu 370
1133 1084 TCCTCGTCCTGGCGACATGACCTGCGCGCCCTGCGTCATCGCATGCC 1133
1133 371 GlnAlaGluProAlaGluSerArgTyrMetAspIleIleTyrThrSe 387
1133 1134 AGCGGCGCTCTCTCATCTGCGACCCCTGCTGCTGCTGCTTGCACAG 1183
1133 387 rcGlyPheLeuAlaValAlaMetAlaIleValIleValLeuGlySargM 404
1184 CCCAGAGAGACCGGTGACACCCCGCGCTGCCCTCC 1221
404 etGlnThrProHisSerLysGlnThrLeuGlnProPro 416

seq_name: sp_rodent:Q9QVY7

seq_documentation_block:
ID Q9QVY7 PRELIMINARY; PRT; 822 AA.
AC Q9QVY7;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE ACIDIC FIBROBLAST GROWTH FACTOR RECEPTOR ISOFORM B, AFGR-B.
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94375484; Pubmed=8089146;
RX Takagi Y, Shrivastava S., Miki T., Sakaguchi K.;
RT "Molecular cloning and expression of the acidic fibroblast growth
RT factor receptors in a rat parathyroid cell line (PT-r). Parathyroid
RT cell-specific calcium-dependent change of ligand accessibility and
RT covalent attachment of heparan sulfate glycosaminoglycan to the
RT receptors."
RL J Biol. Chem. 269:23743-23749(1994).
RU HSP; P06239; 3LCK.
DR InterPro: IPR000719; -.
DR InterPro: IPR001245; -.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 1g; 3.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR SMART: SM00219; TyrcK; 1.
KW ATP-binding; transferase; Tyrosine-protein kinase.
SQ SEQUENCE 822 AA; 92266 MW; 9658E2454023C695 CRC64;

alignment_scores:
Quality: 455.50 Length: 378
Ratio: 1.947 Gaps: 9
Percent Similarity: 61.905 Percent Identity: 31.217

Alignment_block:
US-09-598-042A-2 x Q9QVY7 ..
Align seg 1/1 to: Q9QVY7 from: 1 to: 822

112 GCGCGCATATGCGGCTCAGTGCCTGAGAGGGGAGCCGCCCGCCT 161
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
55 GlyIuSerLeuGluLeuArgCysMetLeuLys.....AspAlaIaVala 69
162 GACCATGTGAGCAAGATGGCGGACCATCCACAGCGGCTGAGACCGCT 211
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
69 lIleSerTrpThrLysAspGly.....ValHisLeuGlyProAsnAsna 84
212 TCCGGGTCTG...CCGAGGGGCTGAAGTGAACAGTGTGAGCGGGAG 258
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

84 rgThrValLeuIleGlyIuTyrLeuGlnIleLysGlyAlaThrProArg 100
259 GATCGCGGCTGTACGTGAGAGGCCAACAGCGCTTCGGACGCTGAG 308
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
101 AspSerGlyLeuTyrAlaCysAlaIaIaIaArgThrValAspSerGluTh 117
309 CGTCAACATACACCCCTGCTGCTGATGATGATGATGATGATGATGATG 358
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
117 rLeuTyrPheMetValAsnValThrAspAlaIleSerSerGlyAspArg 134
359 GCTTGGGCGCGGACAGCTCTCTGCGGGGTCAAGAGAGACCGCCACGAG 408
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
134 lAspAspThrAspThrSerGlnAspPheValSerGlnAsnArgSerAsn 150
409 CAGTGGGACAGACCGCGCTTCACACAGCCCTCCAGATGATGAGCGCGGT 458
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 GlnArgAla...ProTyrTrpThrAsnThrGlnLysMetGlnLysArgLe 166
459 GATCGCAGCGCGCGGAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTG 508
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
166 uHisAlaValAlProAlaIaIaAsnThrValLysPheArgCysProAlaGly 183
509 GGCACCCCTCGCGCCGACATGACATGATGATGATGATGATGATGATG 558
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
183 LysAsnProThrProThrMetArgTyrLeuLysAsnGlyLysGluPheLys 199
559 .....CGCCAGAGGCGCGCTGAGCCCGAGAGAGAGAGAGAGAGACT 599
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
200 GlnGlnHisArgIleGlyGlyTyrLysValArgAsnGlnHisTrpSerLe 216
600 GAGCTCGAAGAACCTGCGGCGGAGGAGCGGCAATGATGATGATGCTCG 649
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
216 uIleMetGluSerValValProSerAspLysGlyLysGlnTyrThrCysLeu 233
650 TGTCAACCGCGCGGCGGCGGATGATGATGATGATGATGATGATGATG 699
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
233 alGlnAsnGlnIuTyrGlySerIleAsnHisThrTyrHisLeuAspValVal 249
700 CAGCGGACCCGTTCCAGCGCGCTGCTGATGATGATGATGATGATGATG 749
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
250 GluArgSerProHisArgProIleLeuGlnAlaGlyLeuProAlaAsnAl 266
750 GACGCTGAGCTTCGGGCGGACCATGCTCTCCAGTGAAGTGGCCAGCG 799
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
266 aserThrValValGlyLysAspValGluPheValLysLysAlaTyrSera 283
800 ACGTGAAGCGCGGTGATCCAGTGCCTGATGATGATGATGATGATGATG 837
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
283 sPAlaGlnProHisIleGlnTrpIleLysHisValGlnLysAsnGlySer 299
838 ...TACGGCGCGGAGGCGCGCCACAACTCCACCATGATGATGATGATG 884
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
300 LysTyrGlyProAspGly..... 305
885 GAAGTTGTGTGCTGCTGCCAGCGGTGACGTG.....TGGTGGCGCGCG 928
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
306 .....LeuProTyrLeuLysValLeuLysHisSerGlyIleAla 318
929 ACGGCTCTACCTCAATAGCTGCTCATGACCGGCTGCGCGGACGAT 978
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
318 snSerSerAsnAlaGluValLeuAlaIaLeuPheAsnValThrIuMetAsp 334
979 GCGGCAATGATACATCTGCTTGGCGGCAACACATGAGGCTACAGTTCG 1028
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
335 AlaGlyIuTyrIleCysLysValSerAsnTyrIleGlyAlaIaAsnGl 351
1029 CAGCGCTCTCTCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1078
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
351 nSerAlaTrpLeuThrValLeuProLysGlnAlaPro..... 364
1079 TGGCTCTCTCGCTCGGCGCATGCTGCGCTGCGCTGCGCTGCTGCTGCT 1128
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
365 ..ValArgGlnLysGlnIleThrAlaSerProAspTyrLeuGlnIleAla 380

```







```

|||||..... ||||| ..... |||||..... |||||.....
241 ArgValGlySerLeuThrTyrThrPheLeuSprValLeuGluArgSe 257
708 CCGTTTCCAAAGCCGCTGTCACAGGACGACCCCGTGAACAGACGCTGG 757
: |||||..... |||||..... |||||..... |||||.....
257 rSerHisArgProIleLeuGlnAlaGlyLeuProIleAsnThrThrAla 274
758 ACTTCGGGGGACCAAGCTCTTCCAGTGCAGAGTGCAGGACGACGTAG 807
|||||..... |||||..... |||||..... |||||.....
274 rValGlySerAspValGluPheTyrCysValTyrSerAspAlaGln 290
808 CCGGTATCCAGTGCAGGCGGTGAGTACGCGCGCGGAGGGCGCGCA 857
||| |||||..... |||||..... |||||..... |||||.....
291 ProIleIleGlnThrLeuLysHisIleGlu..... 300
858 CAACTCCACATCGATGTGGCGGCGCAGAACTTT..... 891
|||||..... |||||..... |||||..... |||||.....
301 .....ValAsnGlySerArgPheGlyProAspArgPheP 312
892 .....GTGCTGCTGCCACGGGTGACGTGTGCTGCGGCCGACGCGC 933
|||||..... |||||..... |||||..... |||||.....
312 rGlyValGlnValLeuLysThrAlaAspIle.....AsnThr 324
934 TCCAACTCATATAGCTGTATACACCCGCTGCCGCGCAGAGATGCGGG 983
||| ..... |||||..... |||||..... |||||.....
325 SerGluValGluValLeuHisLeuArgAsnIleThrMetGluAspAlaG 341
984 CATGTACATGTGCTGTGGCGCAACACCATGAGGTACAGTCTCCAGCG 1033
||| |||||..... |||||..... |||||..... |||||.....
341 yGluTyrThrCysLeuAlaGlyAsnSerIleGlyLeuSerHisGlnSer 358
1034 CTTTCCTGACCGCTGCTGCCAGAACCAACGCGCAGGCGCACTGTGGCC 1083
|||||..... |||||..... |||||..... |||||.....
358 IatPheLeuThrValLeuSerAsnGlu.....AspPheLeuGlu 370
1084 TCCGCTGCTGCTGCGCAATGACCTGCGCGCGCGCGGTATCGGATGCC 1133
|||||..... |||||..... |||||..... |||||.....
371 GlnAlaGluProAlaGluSerArgTyrMetAspIleIleIleTyrThrSe 387
1134 AGCGGCGCTGTCTTCATCTGTGGACCCCTGCTGCTGTGCTTGGCCAG 1183
: |||||..... |||||..... |||||..... |||||.....
387 rGlyPheLeuAlaValAlaMetAlaIleValIleValLeuGlySerArg 404
1184 CCCAGAGAGACCGGTGACCCCGCGCGCTGCCCGCTCCC 1221
||| |||||..... |||||..... |||||..... |||||.....
404 etGlnThrProHisSerLysGlnThrLeuGlnProPro 416
seq_name: sp_veterebrate:Q9PSC9
seq_documentation_block:
ID Q9PSC9 PRELIMINARY; PRT; 480 AA.
AC Q9PSC9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE FIBROBLAST GROWTH FACTOR RECEPTOR 2 ISOFORM IIIC ALPHA, XFGFR2 IIIC
DE ALPHA.
OS Xenopus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
OC Xenopodidae.
OX NCBI_TaxID=8353;
RN (1)
RP SEQUENCE FROM N.A.
RP MEDLINE=95204476; PubMed=7896824;
RA Mathieu M., Kiefer P., Mason I., Dickson C.;
RT "Fibroblast growth factor (FGF) 3 from Xenopus laevis (XFGF3) binds
RT with high affinity to FGF receptor 2."
RL J. Biol. Chem. 270:6779-6787(1995).
DR InterPro: IPR003006;
DR Pfam: PF00047; Ig: 3.
DR SMART: SM00408; IgC2; 1.

```

```

SQ SEQUENCE 480 AA; 53275 MW; C8C3AB636D819A75 CRC64;
alignment_scores:
Quality: 453.00 Length: 409
Ratio: 1.880 Gaps: 13
Percent Similarity: 58.924 Percent Identity: 30.073
alignment_block:
US-09-598-042a-2 x Q9PSC9 ..
Align seg 1/1 to: Q9PSC9 from: 1 to: 480
112 GCGCCGCACTATGCGGCTGTCAGTGCACAGTGAAGGGGAGACCGCGCGCT 161
|||||..... |||||..... |||||..... |||||.....
71 GlyIleuProIleAspLeuArgCysProIleAlaAspGly.....ProIle 85
162 GACCATGTGGACCAAGATGGCGGCACCATCCACAGCGGCTGAGCCGCT 211
||| |||||..... |||||..... |||||..... |||||.....
85 uValThrThrPheLysAspGlyAlaLysLeuGluValAsn.....AsnArgT 101
212 TCCGCGCTGCGCGGAGGCTGAAGTGAAGCAAGTGAAGCGGAGAGAT 261
|||||..... |||||..... |||||..... |||||.....
101 hrLeuIleValArgThrTyrLeuGlnIleLysGluSerThrTyrArgAsp 117
262 GCGGCGGTGTACGTGTGCAAGGCGCACAAGGCTTTCGCGACGCTGAGCT 311
|||||..... |||||..... |||||..... |||||.....
118 SerGlyLeuTyrAlaCysSerValLeuLysAsnSerHisPhePheHisVa 134
312 CAAGTACACCCCTGCTGCTGATGATGACATTAGCCGAGGAAGAGAGCC 361
|||||..... |||||..... |||||..... |||||.....
134 IAsnValThr..... 137
362 TGGGCGCGCAGACTCTCTGTGGGGGTCAAGAGAC.....CGCGCGAC 405
|||||..... |||||..... |||||..... |||||.....
138 .....GluAlaSerSerGergLysAspAspGluAspAspAspLysSer 152
406 CAGCAGTGGGCGACA.....CGCGCTTCACACA 434
|||||..... |||||..... |||||..... |||||.....
153 GlnAspPheThrAsnAspAsnAlaIleArgAlaProTyrTrpThrAs 169
435 GCGCTCCAAAGATGAGCGCGCGGTGATCGCACGCGCGCGGTGACTCGG 484
: |||||..... |||||..... |||||..... |||||.....
169 nThrIuLysMetGluLysLysLeuHisAlaValProAlaAlaAsnThrV 186
485 TCGCGCTCAAGTGCCTGCGCAGCGGCGACCTCGCGCGCATCAGTGG 534
|||||..... |||||..... |||||..... |||||.....
186 alLysLeuArgCysProAlaIleGlyLysAsnProThrProArgMetArgTrp 202
535 ATGAAGGACGACCGACCTTGACG.....CGCGGAGAGCGCGCTGA 575
|||||..... |||||..... |||||..... |||||.....
203 LeuLysAsnGlyLysGluPheLysGlnGlnHisArgIleGlyTyr 219
576 GCGCAGGAAGAAGTGAAGTGAAGCTGAAGAACCTGCGCGCGGAG 625
: |||||..... |||||..... |||||..... |||||.....
219 sValArgAsnGlnHisTyrSerLeuIleMetGluSerValAlaProSer 236
626 ACAGGCGCAATATACACTGCGCGGTGTGCAACCGCGCGGCGCGCATCAAC 675
|||||..... |||||..... |||||..... |||||.....
236 sPlyGlyIleTyrThrCysIleValGluAsnGlnHisGlySerIleAsn 252
676 GCGACCTACAAAGTGATGTGATCCAGGCGACCGCTTCACAGCGCGTGT 725
|||||..... |||||..... |||||..... |||||.....
253 HisThrTyrHisLeuAspValIleGluArgSerHisArgProIle 269
726 CACAGGCGACGACCGCGTGAACACAGCGTGAAGTTCGGGGGAGACAGT 775
||| ..... |||||..... |||||..... |||||.....
269 uGlnAlaGlyLeuProAlaAsnThrThrAlaValAlaGlyLysAspAlaG 286
776 CTTTCAGTGCAGAGTGCAGCGCAGTGAAGCGGTGATCAGTGGCTG 825
|||||..... |||||..... |||||..... |||||.....
286 IuPheValCysLysValTyrSerAspAlaGlnProHisIleArgTyrVal 302

```





```

276 heValcysLysValItyrSeraspAlaInProHisIleGlnTrpIleLys 292
829 CGCGTGGAG.....TACGGCCCGAGGGCCCGCCACACTC 863
293 HisValGluLysAsnGlySerLysItyrGlyProaspGly..... 305
864 CACCATGATGTGGGGCGGCAGAGTTGTGTGTCGCCCGGAGGTGACG 913
306 .....LeuProTyrIleuLysV 311
914 TGTGTGCGGGCGCCGAGCGCTCTACTCAATPAAG.....CTG 951
311 allLeuLysAlaIleGlyValAsnThrThrAspLysGluIleGluValLeu 327
952 CTATTCACCCCTGCCCGCCAGAGATGCGGCGATGTACATCTGCTTGG 1001
328 TyrIleLeuArgAsnValIthrPheGluAspAlaGlyIuTyrThrCysLeuAl 344
1002 CGCCACACACATGGGCTACAGCTTCGGCAGCGCTTCCTCCACCGTGTGC 1051
344 agLYAsnSerIleGlyIleSerPheHisSerAlaTrpLeuThrValLeuP 361
1052 CAGACCCAAAACGCCAGGGCCACCTGTGGCCCTCTGCTCGGCCACT 1101
361 roAlaPro.....ValArgGluLysGluIleThr 370
1102 AGCTGCGCGTGGCCCGTGTATCGCATCGCATCCAGCGCGCTGTCTTCA 1151
371 AlAserProAspTyrLeuGluIleAlaIleTyrCysIleGlyValPheLe 387
1152 CCTGGAGCAGCTGCTGCTGTGGCTTTCGAGGCCAGAAAGAGCGGTGCA 1201
387 uileValCysMetValIleThrValIleValCysArgMetLysThrThrS 404
1202 CCCCCGCGCCT 1212
404 erLysLysPro 407
seq_name: sp_human:090IH3
seq_documentation_block:
ID 090IH3 PRELIMINARY; PRT; 768 AA.
AC 090IH3:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE K-SAM-IIO3.
GN K-SAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RA Ueda T., Sasaki H., Kuwahara Y., Nezu M., Shibuya T., Sakamoto H.,
RA Ishii H., Yanagihara K., Matune K., Makunuchi M., Terada M.;
RT "Deletion of the carboxyl-terminal exons of k-sam/EGFR2 by short
RT homology-mediated recombination, generating preferential expression of
RT specific mRNAs.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
CC EMBL; AB030078; BAA89301.1; -.
CC HSSP; P12931; 1FMK.
DR InterPro: IPR000719; -.
DR InterPro: IPR001245; -.
DR InterPro: IPR003006; -.
DR Pfam; PF00047; 1g; 3.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

```

```

DR SMART; SM00219; TyKc; 1.
KW ATP-binding; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 768 AA; 86208 MW; A6F1A35DD69FAAADA CRC64;

alignment_scores:
Quality: 452.50 Length: 385
Ratio: 1.926 Gaps: 10
Percent Similarity: 61.039 Percent Identity: 30.909

alignment_block:
US-09-598-042A-2 x 090IH3 ..
Align seg 1/1 to: 090IH3 from: 1 to: 768

100 GTGGCCCGCGTGGCCGACATAGCGGCTGAGTCCCGAGAGGGGGA 149
||||| |||||.....
51 ValAlaIleProGlyGluSerLeuGluValAlaGlySerLeuLys..... 65
150 CCGCGCGCGCTGACCATGTGACCAAGATGGCCGACCATCCACAGC 199
||| .....
66 AspAlaIleValIleSerTrpThrLysAspGly.....ValHisLeuG 80
200 GGTGGAGCGCGCTCCGCGTGCCTG...CCGACAGGCGTGAAGTGAAGC 246
|| .....
80 LyrProAsnAsnArgThrValLeuIleGlyIuTyrLeuGlnIleLysGly 96
247 GTGGAGCGGAGATGCCGCGGTGTACGTGTGCAAGGCCACCAAGCGCT 296
|| .....
97 AlaThrProArgAspSerGlyLeuTyrAlaCysThrAlaSerArgThrVa 113
297 CGGCAGCTGAGCGTCACTACCCCTGCTGTGTGATGATGATAC 346
||| .....
113 LAspSerGluThrTrpTyrPheMetValAsnValThrAspAlaIleSerS 130
347 CAGGAAGAGAGCCTGGGCGCCGACAGCTCTGCGGGGTCAAGAGAGC 396
||| .....
130 erGlyAspAsp.....GluAspAspThrAspGlyAlaGluAsp 142
397 CCGCGCAGCCAG.....CAGTGGGACAGACCGCGCTTCACACAGCC 437
||| .....
143 PheValSerGluAsnSerAsnAsnLysArgAlaProTyrTrpThrAsnThr 159
438 CTCACAGATGAGCGCGCGGTGATGCGACGGCCGCTGGTACTCGTGC 487
||| .....
159 rGluLysMetGluLysArgLeuHisAlaValProAlaIleAsnThrValL 176
488 GGTCAAGTGGTGGCGCCAGCGGCGCCTCGCGCCGACATGACGTGATG 537
||| .....
176 yspHeArgCysProAlaGlyLysAsnProMetProThrMetArgTrpLeu 192
538 AAGCAGCAGCAGCGCTTGACG.....CGCCAGAGCGCGGTGAGCC 578
||| .....
193 LysAsnGlyLysGluPheLysGlnIleHisArgIleGlyTyrLysVa 209
579 CAGGAAGAAGAGTGGACATGAGCCTGAAGAACCTGCGCGGAGAGACA 628
||| .....
209 LArgAsnGlnHisTrpSerLeuIleMetGluSerValValProSerAspL 226
629 GCGGCAATACACCTGCGCGGTGTGGAACCGCGCGGCGCATCAAGCC 678
||| .....
226 ysgLYAsnTrpThrCysValAlaGluAsnGluTyrGlySerIleAsnHis 242
679 ACCTACAGGTGATGTATCAGCGAGCGCGTTCACAGCCCGTGTCTAC 728
||||| .....
243 ThrTyrHisLeuAspValValGluArgSerProHisArgProIleLeuGl 259
729 AGGACGACCCCGTGAACACAGCGGTGACTCGGGGGGACACAGTCT 778
||| .....
259 nAlaGlyLeuProAlaAsnAlaSerThrValValGlyLysAspValGluP 276
779 TCcATGTCAGAGTGGCAGCGACGTAAGACCGGTGATTCATGCGTGAAG 828
||| .....

```

```

276 heValcysLysValTyrSerAspAlaGlnProHisIleGlnTrpIleLeu 292
829 CCGGTGAG.....TAGGGCCGAGGGCCGCCCAACTC 863
293 HisValGlnLysAsnGlySerLysTyrGlyProAspGly..... 305
864 CACCATCATGTGGGGCCGAGAGTTTGTGTGTCGTCGACCGGTTGAC 913
306 .....LeuProTyrLeuLys 311
914 TG.....TGTCGCGCCGACGCTCTACCTCAATAAGCTGTATC 957
311 allLeuLysHisSerGlyIleAsnSerSerAsnAlaGlnValIleAlaLeu 327
958 ACCGGTCCCGCCGAGCATGGCGGATGTACATCTGCTGTCGCGCA 1007
328 PheAsnValThrGlnAlaAsnGlnSerAlaTrpLeuThrValLeuProLys 361
1008 CACCATGGCTACAGCTTCGACGCGCTTCCTCACCCTGTGTCGACACC 1057
344 nTyrIleGlyGlnAlaAsnGlnSerAlaTrpLeuThrValLeuProLys 361
1058 CAATACCGCCGAGCCACTGTGGCTCTCTGCTGTCGCGCACTAGCTG 1107
361 LngIleAlaProGly.....ArgGlnLysGlnIleThrAlaSer 373
1108 CCGTGGCCGCTGTGTCATCGGATCCGAGCGGCTGTCTCATCGGG 1157
374 ProAspTyrLeuGlnIleAlaIleTyrCysIleGlyValPheLeuIleAl 390
1158 CACCTGCTCTCTGTGGCTTCGACGCGCCAGAGAACCGCTGCACCCCG 1207
390 acyMetValValThrValIleLeuLysArgMetLysAsnThrThrLysL 407
1208 CCGCT 1212
407 yspPro 408
seq_name: sp_human:Q9UIH5
seq_documentation_block:
ID Q9UIH5 PRELIMINARY; PRT; 771 AA.
AC Q9UIH5;
DR 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE K-SAM-1101.
GN K-SAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ueda T., Sasaghi H., Kuwahara Y., Nezu M., Shibuya T., Sakamoto H.,
RA Ishii H., Yanagihara K., Matune K., Makuchi M., Terada M.;
RT "Deletion of the carboxyl-terminal exons of K-sam/GFR2 by short
RT homology-mediated recombination, generating preferential expression of
RT specific mRNAs."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL; AB030076; BAAB9299.1; -.
DR HSP; P12931; IFMK.
DR InterPro; IPR000719; -.
DR InterPro; IPR001245; -.
DR InterPro; IPR003006; -.
DR Pfam; PF00047; 1g; 3.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PRO0109; TYRKINASE.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

```

```

DR SMART; SM00219; TYRC; 1.
KW ATP-binding; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 771 AA; 86505 MW; BRC73746F1A5DD6 CMC64;

alignment_scores:
Quality: 452.50 Length: 385
Ratio: 1.926 Gaps: 10
Percent Similarity: 61.039 Percent Identity: 30.909

alignment_block:
US-09-598-042a-2 x Q9UIH5 ..
Align seg 1/1 to: Q9UIH5 from: 1 to: 771

100 GTGCGCCGCGTGGCGCCGACATAGCGGCTGCACTGCGCCATGAGGGGCA 149
||||| |||.....|
51 ValAlaAlaProGlyIleSerLeuGlnValArgCysLeuLys..... 65
150 CCGCGCCGCGCTGACCATGTGACCAAGGATGCGCCGACCATCCAGCGG 199
||| .....|
66 AspAlaAlaValIleSerTrpThrLysAspGly.....ValHisLeu 80
200 GCTGAGCGCGCTTCGCGTCTG...CCGAGGGGCTGAGGTGAGGACG 246
|| .....|
80 LysAsnAsnArgThrValLeuIleGlyLutyrLeuGlnIleLysGly 96
247 GTGAGCGGGAGATGCGCGGCTGTACGTGTGCAAGCCACCAACGGCTT 296
||| .....|
97 AlaThrProArgAspSerGlyLeuTyrAlaCysThrAlaSerArgThy 113
297 CCGGACCTGTGAGCTGACACTACACCTCGTCTGTGATGATGATGAC 346
|| .....|
113 LysSerGlnThrTrpTyrPheMetValAsnValThrAspAlaIleSer 130
347 CAGGGAAGGAGAGCTGGGGCCCGACAGCTCTGTGGGGTCAAGAGGAC 396
|| .....|
130 erGlyAspAsp.....GluAspAspThrAspGlyAlaGluAsp 142
397 CCGCGCCAGCG...CAGTGGGACAGCGCGGCTGCACAGCG 437
||| .....|
143 PheValSerGlnAsnSerAsnAsnLysArgAlaProLutyrThrAsn 159
438 CTCGAGATGAGGCGCGGCGGATGCGACGCGCGCTGTGATGCTGCG 487
||| .....|
159 rGlnLysMetGlnLysArgLeuHisAlaValProAlaAlaAsnThrVal 176
488 GGCTCAAGTGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 537
||| .....|
176 yspPheArgCysProAlaGlyLysAsnProMetProThrMetArgTrpLeu 192
538 AAGGAGACCGAGCGCTTGACG.....CGCCAGAGCGCGCTGAGCC 578
||| .....|
193 LysAsnGlyLysGlnPheLysGlnGlnHisArgIleGlyLutyrLys 209
579 CAGGAAGAAGACTGAGACACTGAGCTGAAGACTGCGCGCGGAGACA 628
||| .....|
209 LysGlnGlnHisTrpSerLeuIleMetGlnSerValProSerAsp 226
629 GCGGCAATACCTGCGCGGCTGCAAGCGCGGCGGCGGCGGCGGCGG 678
||| .....|
226 yspGlyAsnTyrThrCysValValGlnAsnGlnLutyrIleSerIleAsnHis 242
679 ACCTACAGGTGATGATCCAGCGGACCGCTGCAAGCGGCTGCTGAC 728
||||| .....|
243 ThrTyrHisLeuAspValValGlnArgSerProHisArgProIleLeu 259
729 AGGACGACCGCCGTAACAGCGGCTGACCTTGGGGGAGGACGAGCTCT 778
||| .....|
259 nAlaGlyLeuProAlaAsnAlaSerThrValValGlyLysAspValGln 276
779 TCCAGTGCAGAGTGGCGAGGAGCGTGAAGCGGCTGATCCAGTGGCTGAG 828
|| |.....|

```

```

276 heValcysLysValItyrSeraspAlaInProHsIleGlnTrpIleLys 292
829 CGCGTGGAG.....TACGGCCCGAGGGCCCGCACACTC 863
      :|||||
293 HisValGluYsaNglySerItyrGlyProAspGly..... 305
864 CACCATGATGTGGCGGCGAGAGTTGTGTGTCGCCCGGCGTACG 913
306 .....LeuProItyrLeuLysV 311
914 TG.....TGTCGGCGCCGAGCGCTCTACCTCAATACCTGCTCATC 957
      ||| :|||
311 alleuLysHisSerGlyIleAsnSerSerAsnAlaGluValLeuAlaLeu 327
958 ACCGTCGCCCGCCAGACGATCGCGCATGTACATCTGCTGGCGCCA 1007
      :||| :|||
328 pheAsnValThrGluAlaAspAlaGlyIuItyrIleCysLysValSerAs 344
1008 CACCATGGGCTACAGCTTCGCGAGCGCTCTCTACCGCTGCTGCAGACC 1057
      :||| :|||
344 nTyriIleGlyInAlaAsnGlnSerAlaTrpLeuThrValLeuProLysG 361
1058 CAACACCGCCAGCGCCACCTGTGCTCTCTCTGCTGCGCCACTAGCCTG 1107
      :||| :|||
361 lngInAlaProGly.....ArgGluLysGluIleThrAlaSer 373
1108 CCGTCGGCCCGTGTCAATCGGCATCCACCGCGCGCTGTTCATCTGGG 1157
      :||| :|||
374 ProAspItyrLeuGluIleAlaIleItyrCysIleGlyValPheLeuIleAl 390
1158 CACCGCTGCTCTGTCGGCTTTCGCGCCGACCAAGACCGCTGCACCCCG 1207
      :||| :|||
390 aCysMetValValThrValIleLeuCysArgMetLysAsnThrItyrSL 407
1208 CGCCT 1212
      |||
407 yspPro 408

seq_name: sp_human:09UTH4
seq_documentation_block:
ID 09UTH4 PRELIMINARY; PRT; 817 AA.
AC 09UTH4:
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE K-SAM-II02.
GN K-SAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ueda T., Sasaki H., Kuwahara Y., Nezu M., Shibuya T., Sakamoto H.,
RA Ishii H., Yanagihara K., Matune K., Makunuchi M., Terada M.;
RT "Deletion of the carboxyl-terminal exons of K-sam/FeFR2 by short
RT homology-mediated recombination, generating preferential expression of
RT specific mRNAs."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL; AB030077; BAA89300.1; -.
DR HSSP; P12931; 1FMK.
DR InterPro; IPR000719; -.
DR InterPro; IPR001245; -.
DR InterPro; IPR003006; -.
DR Pfam; PF00047; 1g; 3.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PRO0109; TYRKINASE.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

```

```

DR SMART; SM00219; TYRK; 1.
KW ATP-binding; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 817 AA; 91718 MW; A84ED3EFBD0D1877 CRC64;

alignment_scores:
      Quality: 452.50      Length: 385
      Ratio: 1.926      Gaps: 10
Percent Similarity: 61.039      Percent Identity: 30.909

alignment_block:
US-09-598-042A-2 x 09UTH4
Align seg 1/1 to: 09UTH4 from: 1 to: 817

100 GTGGCCCGCGTGGCGCGCATAGCGGCTGACAGTCCGAGAGGGGA 149
      |||||
51 ValAlaAlaProGlyGluSerLeuGluValArgCysLeuLys..... 65
150 CCGCGCGCGCTGACCATGTGGACCAAGATGGCGCGCACATCCACAGC 199
      :||| :|||
66 .AspAlaAlaValIleSerTrpTrpLysAspDly.....ValHisLeuG 80
200 GCTGGAGCGCGTCCGCGTGGTGG.....CCGAGGCGCTGAGGTGAGCAG 246
      ||| :|||
80 lProAsnAsnArgThrValLeuIleGlyGlyItyrLeuGlnIleLysGly 96
247 GTGGAGCGGAGATGCCGCGCTGTACCTGTGCAAGCGCCACCAACGGCT 296
      :||| :|||
97 AlaThrProArgAspSerClyLeuItyrAlaCysTrpAlaSerArgThrVa 113
297 CGGACCGCTGACGTCACATACACCTCTGCTGCTGATGACATTAAGC 346
      :||| :|||
113 lAspSerGluThrTrpItyrPheMetValAsnValThrAspAlaIleSerS 130
347 CAGGGAAGACAGACGCTGGCGCCGACAGCTCTGCGGGGTGAGAGAGAC 396
      |||
130 erGlyAspAsp.....GluAspAspThrAspDlyAlaGluAsp 142
397 CCGCGCACCCAG.....CAGTGGCGACAGCCGCGCTTCACACAGCC 437
      :||| :|||
143 PheValSerGluAsnSerAsnAsnLysArgAlaProGlyTrpThrAsnTh 159
438 STCCAAAGATGAGCGCGCGTATGCGACGCGCGCTGATGATCCGCTGC 487
      :||| :|||
159 rGluLysMetGluLysArgLysHisAlaValProAlaAlaAsnThrValL 176
488 GGCACAACTGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 537
      :||| :|||
176 yspPheArgCysProAlaGlyIleLysnProMetProThrMetArgTrpLeu 192
538 AAGGACGACAGCGCTTGACG.....CGCCAGAGCGCGCGTACGCC 578
      :||| :|||
193 LysAsnGlyLysGluPheLysGlnGlnHisArgIleGlyGlyItyrLysVa 209
579 CAGGAAGAAGAGTGGACACTGAGCGCTGAGAACCTGCGCGGAGAGACA 628
      :||| :|||
209 lArgAsnGlnHisTrpSerLeuIleMetGluSerValValProSerAspRL 226
629 CGCGCAATACACTCGCGCGCTGTGCGACCGCGCGCGCGCGCGCGCGCG 678
      :||| :|||
226 ysgLysnTrpThrCysValAlaGluAsnGlnItyrGlySerIleHis 242
679 ACCTCAAGATGTGATGTATCCAGCGGCGCGCTTCCAGCGCGCTGTCTAC 728
      |||||
243 ThrItyrHisLeuAspValValGluArgSerProHisArgProIleLeuG 259
729 AGGACGACCGCGCTGAAACAGCAGCGTGGACTTGGGGGGAACAAGTCT 778
      :||| :|||
259 nAlaGlyLeuProAlaAsnAlaSerThrValValGlyLysAspValGluP 276
779 TCCAGTGAAGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 828
      |||

```

```

276 hevalcysylvaltyrserasplalaglprohislleclntrilelys 292
829 CCGGTGGAG.....TACGGCCGAGGGCCGACACTC 863
      :::::|||||
293 HlsvaIglulysanllyserlytyrlyproaspIy..... 305
864 CACCATCATGTGGGGGGCCAGAGTTTGTGTCGTCGACGGGTGAG 913
306 .....LeuProTyrLeuLysV 311
914 TG.....TGTCGGCGCCGACGGCTCTACCTCATATAAGCTGTATC 957
      ||| :::::|||||
311 allenyshtserserlyleasnsersasnalaglvalleualaleu 327
958 ACCCGTGGCCGAGACATGTATCATCTGCTTGGCCCA 1007
      :::::|||||
328 pheasvalthrcIuIalaspIaglygltyrIleCysylvalseras 344
1008 CACCATGGCTACAGCTCCGACGCTCTCCACGCTGTCGACAGCC 1057
      ||| :::::|||||
344 nlyrIleIglIalaspIaglnseralatrpleuthrvalleuprollys 361
374 proaspIytleuIlealaleItyrCysIleIglvalpheuIleal 390
1158 CACCTGCTCTGCTGCTTGCAGGCGCAGAGACCGGTGACCCCG 1207
      : :::::|||||
390 acysmetvalvalthValIleucysarIgmetylsasnthrIlelys 407
1208 CGCCT 1212
      |||
407 yspro 408
seq_name: sp_human:Q9UIH8
seq_documentation_block:
ID Q9UIH8 PRELIMINARY: PRT: 819 AA.
AC Q9UIH8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE K-SAM-TIH1.
GN K-SAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ueda T., Sasaki H., Kuwahara Y., Nezu M., Shibuya T., Sakamoto H.,
RA Ishii H., Yanagihara K., Maifune K., Makuchi M., Terada M.;
RT "Deletion of the carboxyl-terminal exons of K-sam/EGFR2 by short
RT homology-mediated recombination, generating preferential expression of
RT specific mRNAs."
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
EMBL: AB030073; BA089296.1; -
DR HSPB; P12931; 1PK.
DR InterPro: IPR000719; -
DR InterPro: IPR001245; -
DR InterPro: IPR003006; -
DR Pfam: PF00047; 1g; 3.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.

```

```

DR SMART: SM00219; TYRKC; 1.
KW AMP-binding; Transferase; Tyrosine-protein kinase.
SO SEQUENCE 819 AA; 91664 MW; D2AAE6A6ABA6E CRC64;

alignment_scores:
  Quality: 452.50      Length: 385
  Ratio: 1.926        Gaps: 10
  Percent Similarity: 61.039      Percent Identity: 30.909

alignment_block:
US-09-598-042a-2 x Q9UIH8
Align seg 1/1 to: Q9UIH8 from: 1 to: 819

100 GTGGCCCGGCTGGCCGACATATGGCGCTGCAGTGGCCGACGAGGGGGA 149
      ||||| |||||
51 ValAlaIalProglYglInsersleuGlualArgCysleuLys..... 65
150 CCGCGCCGCTGACCATGTGACCAAGAGATGCGGCACACATCCAGCG 199
      :::::|||||
66 AsplalalavalIlesertrIphIlyspsIy.....ValHlsleuG 80
200 GCTGAGCGCTTCGCGCTG...CGCAGGGCTGAAGTGAAGCAG 246
      ||| :::::|||||
80 lYprohnsasnaIgrthValleuIleIglulutyrlleuGlInleIysglY 96
247 GTGAGCGGAGATCCGCGGTACGTTGTCAGGCGCACAGCGCTT 296
      :::::|||||
97 AlathrProIgrAspserselYleuIyralaCysIthralaserIgrthya 113
297 CGGACCGCTGAGCGTCACTACACCTCTGCTGTGATGATTAAGCC 346
      ||| :::::|||||
113 lAspsersglutIhtrIptYrPhemeValasnlValThrAsplalIlesers 130
347 CAGGGAAGAGAGCTGGGCGCCGACAGCTCTCTGGGGTCAAGAGCAG 396
      ||| :::::|||||
130 erclYaspsp.....GluspspsIhtrAsplalYalaglualasp 142
397 CCGCGCCAGCAG.....CAGTGGCAGCAGCGCGCTTCACACAGCC 437
      :::::|||||
143 phevalSerGlInsersasnsnlysalgalalProIytrIphIrsnth 159
438 CTCCAAGATGAGCGCGGGGTATCGACGCGCGGTGATGCTCGTGC 487
      :::::|||||
159 rglulysmetglulysargleuIhisalavalProalalalasnthrValI 176
488 GGTCAAGTGCCTGGCGCAGCGGACCTTCGCGCCACATCACTGATG 537
      :::::|||||
176 yspheargCysProalaglYglYasnPromerProthmetatgrtpleu 192
538 AAGGAGACAGCGCTTGACG.....CGCCAGAGCGCGCTGAGCC 578
      ||||| :::::|||||
193 lYsasnglYlYsgIurpHeIysglIngluIhisArgIleelYlyrIlysva 209
579 CAGGAAGAAGACTGACATGACCTGAAGACCTGCGCGCGAGGAGACA 628
      ||||| :::::|||||
209 lArgasnglnhIstrserIleuIlemetgluserIalValProserAspl 226
629 GCGGCAATACACTCCGCGCTGCACACCGCGCGGCGGCATCAAGCC 678
      :::::|||||
226 ysgIYasnIytrIthCysValalGlualasnglutyrlYserIleasnhIs 242
679 ACCTACAGTGTGATGCCAGCGGACCGCTTCAGAGCGGTGCTGAC 728
      ||||| :::::|||||
243 ThrIyrlhisleuaspvalalGlualArgserProhIasIgrProIleuGl 259
729 AGGACAGCACCCCGTACACAGAGGTGACTTCGGGGAGCACAGCTCT 778
      :::::|||||
259 nalaglYleuProalalasnIaserthrValalglYlYasvalglulP 276
779 TCAGTGCAGAGTGGCGAGGAGCTGAAGCGGTGATCCAGTGGCTGAAG 828
      ||| |||||

```

```

276 heValcysLysValItyrSeraspAlaInProHisIleGlnTrpIleLys 292
829 CGCGTGGAG.....TACGGCGCCGAGGGCCGCCCACTC 863
      :::::::::::
293 HisValGlnLysAsnGlySerLysTyrGlyProAspGly..... 305
864 CACCATGATGTGGGGCCGACAGATTGTGTGTCCTGCCCAAGGGTGACG 913
306 .....LeuProTyrIleuLys 311
914 TG.....TGTCCGCGCCGACGCGCTCTACCTCAATAAGCTGCTCATC 957
      ||| ::::::::::: ||| ::::
311 allLeuLysHisSerGlyIleAsnSerSerAsnAlaGluValLeuAlaLeu 327
958 ACCCGTGGCCCGCAGACGATCGGGCATGTACATCTGCTGGCGCCCA 1007
      ::::: ::::: ||||||| ||||||| :::::
328 PheAsnValThrGlnAlaAspAlaGlyIleuTyrIleCysLysValSerAs 344
1008 CACCATGGGCTACAGCTCCGACGCGCTCTACACCTGCTGCCAGACC 1057
      | ::::: | ::::: ||||||| |||||||
344 nTyrIleGlyGlnAlaAsnGlnSerAlaTrpLeuThrValLeuProLysG 361
1058 CAACACGCGCAGGGCCACCTGTGGCTCTGCTCGCCCATAGCTG 1107
      ::::: ||||||| :::::
361 GlnGlnAlaProGly.....ArgGlnLysGlnIleThrAlaSer 373
1108 CCGTGGCGCGGTGTCATCGCATCCGACGCGCGCTGTCTTCATCTGGG 1157
      ||| ::::: ||::: ||::: :::::
374 ProAspTyrIleuGlnIleAlaIleTyrCysIleGlyValPheLeuIleAl 390
1158 CACCTGCTCTGTGGCTTGGCCAGGCCCAAGAACGCCGTGCACCCCG 1207
      ::::: ::::: |||
390 acyMetValValThrValIleLeuCysArgMetLysAsnThrTyrLsL 407
1208 CGCCT 1212
      |||
407 yspPro 408

seq_name: sp_human:09U1H7
seq_documentation_block:
ID 09U1H7 PRELIMINARY; PRT: 819 AA.
AC 09U1H7:
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DE 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
DE K-SAM-IIH2 (FRAGMENT).
GN K-SAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ueda T., Sasaki H., Kuwahara Y., Nezu M., Shibuya T., Sakamoto H.,
RA Ishii H., Yanagihara K., Matune K., Makuuchi M., Terada M.,
RA "Deletion of the carboxyl-terminal exons of K-sam/FGFR2 by short
RT homology-mediated recombination, generating preferential expression of
RT specific mRNAs."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL: AB030074; BAA89297.1; -.
DR HSSP: P12931; IFMK.
DR InterPro: IPR000719; -.
DR InterPro: IPR001245; -.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 1g; 3.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.

```

```

DR SMART: SM00219; TYRK; 1.
KW ATP-binding; Transferase; Tyrosine-protein kinase.
FT NON_TER 819 819
SQ SEQUENCE 819 AA; 91739 MW; 52D5955E7B76F4D9 CRC64;

alignment_scores:
      Quality: 452.50      Length: 385
      Ratio: 1.926      Gaps: 10
Percent Similarity: 61.039      Percent Identity: 30.909

alignment_block:
US-09-598-042A-2 x 09U1H7

Align seg 1/1 to: 09U1H7 from: 1 to: 819
100 GTGGCGCGCGTGGCGCCGACATATGCGGTGAGTGCACCGTAGAGGGGA 149
      |||||
51 ValAlaAlaIleProGlyGlnSerLeuGlnValArgCysLeuLeuLys.... 65
150 CCGCGCGCGCGTACCATGTGGACCAAGATGGCCGACCATTCACAGCG 199
      ::::: |||||||
66 AspAlaAlaValIleSerTrpThrLysAspGly.....ValHisLeuG 80
200 GCTGAGCGCGCTCCGCGCTGCTG...CCGACAGGGCTGAGGTGAGAGAG 246
      || ::::: |||||||
80 TyrProAsnAsnArgThrValLeuIleGlyGlyTyrLeuGlnIleLysGly 96
247 GTGGAGCGCGGATGCGCGCTGACCTGTGTCAGGACCGACCGAGCTT 296
      ::::: |||||||
97 AlaThrProAlaGAspSerGlyLeuTyrAlaCysThrAlaSerArgThrVa 113
297 CGGACGCTGAGCGTCAACTACACCTGCTGCTGCTGATGATCATTTAGC 346
      ||| ::::: ||| |||||
113 AspSerGlnThrTrpTyrPheMetValAsnValThrAspAlaIleSerS 130
347 CAGGAAAGAGAGCCTGGGGCCCGACAGCTCTCTGGGGGTCAAGAGAGC 396
      ||| ::::: ||| |||||
130 ergLysAspAsp.....GlnAspAspThrAspGlyAlaGlnLysP 142
397 CCGCGCACCCAG.....CAGTGGCGCAGACCGCGCTTCACACAGCC 437
      ::::: ||::: ||| :::::
143 PheValSerGlnAsnSerAsnAsnLysArgAlaProTyrTrpThrAsnTh 159
438 CTCGAAGATGAGCGCGCGGTGATGCGACGCGCGCTGGGTACTCGTGC 487
      ::::: |||||||
159 rGlnLysMetGlnLysArgLeuHisAlaValProAlaAlaAsnThrValL 176
488 GGCTCAAGTGGCGGCGCAGCGGCGACCTCGCGCCGACATCAGTGTATG 537
      ::::: ||| ::::: ||| :::::
176 yspPheArgCysProAlaGlyGlnAsnProMetProThrMetArgTrpLe 192
538 AAGACAGACAGCGCTTGACG.....CGCCAGAGAGCGCGCTGAGCC 578
      ||||| ::::: ||| :::::
193 LysAsnGlnLysGlnPheLysGlnIleHisArgIleGlyGlyTyrLysVa 209
579 CAGGAAGAAGAGTGGACACTGACCTGAAGAACCTGCGCGCGGAGACA 628
      ||||| ::::: ||| :::::
209 lArgGlnGlnHisTrpSerLeuIleMetGlnSerValValProSerAspL 226
629 GCGGCAATACACCTCGCGCGTGTGGAACCGCGCGCGCGCCATCAAGCC 678
      ::::: |||||||
226 ysgLysAsnTyrThrCysValValGlnAsnGlnTyrGlySerIleAsnHis 242
679 ACCTCAAGGTGATGTATCCAGCGGACCGCTTCAGACCGCGTGTCTAC 728
      ||||| ::::: |||||
243 ThrTyrHisLeuAspValValGlnArgSerProHisArgProIleLeuG 259
729 AGGCACGACCGCGTGAACACGACGAGTGGACTTCGGGGGAGACAGTCT 778
      ::::: |||||
259 nAlaGlyLeuProAlaAsnAlaSerThrValValGlyLysAspValGlnP 276
779 TCAGATGCAAGTGGCGACGCGACGTGAAGCGCGTATCATCGAGTGAAG 828

```



```

276 hevalCysLysValItyrSeraspAlaGlnProHisIleGlnTriPheLys 292
829 CGCGTGGAG.....TACGGCGCCAGGGCCGCCACACTC 863
293 HisValGlnLysAsnGlySerLysTyrGlyProAspGly..... 305
864 CACATCATGTGGGGCGGCAGAACTTGTGGTGGTGGCCGACGGGTGACG 913
306 .....LeuProTyrLeuLysV 311
914 TG.....TGGTCGGCGCCGACGGCTCTACCTCATATAAGCTGCTATC 957
111 .....TACGGCGCCAGGGCCGCCACACTC 863
311 allLeuLysHisSerGlyLysAsnSerSerAsnAlaGlnValLeuAlaLeu 327
958 ACCGTGCGCCGCGACGACGATCGGGCATGTACATCATGTGCTGGCCGCA 1007
328 PheAsnValThrGlnAlaAspAlaGlyLysTyrLysValSerAs 344
1008 CACCATGGGCTACAGCTCCGGAGCGCTCTGCTACGCTGGTGGCCAGAC 1057
344 nTyrIleGlyAlaAlaAsnGlnSerAlaTrpLeuThrValLeuProLysG 361
1058 CAAACCCGACGAGCGACCTGTGGCTCTGCTCTGCGCCACATAGCTG 1107
361 GlnAlaAlaProGly.....ArgGlnLysGlnIleThrAlaSer 373
1108 CCGTGGCGCCGCTGATCGCATCCGACCGCGCTGTCTCATCTGGG 1157
374 ProAspTyrLeuGlnIleAlaIleLysTyrLysIleGlyValPheLeuIleAl 390
1158 CACCGTCTGCTGTGGCTTGGCCAGCGCCAGAAAGCCGTCACCCCG 1207
390 acysMetValValThrValIleLeuCysArgMetLysAsnThrLysL 407
1208 CGCCT 1212
407 yspco 408

seq_name: sp_human:09UIN6
seq_definition_block:
ID 09UIN6 PRELIMINARY; PRT; 830 AA.
AC 09UIN6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE K-SAM-IIH3.
GN K-SAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ueda T., Sasaki H., Kuwahara Y., Nezu M., Shibuya T., Sakamoto H.,
RA Ishii H., Yanagihara K., Matune K., Makuuchi M., Terada M.;
RT Deletion of the carboxyl-terminal exons of K-sam/FGFR2 by short
RT homology-mediated recombination, generating preferential expression of
RT specific mRNAs.
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
CC EMBL: AB030075; BAA89298.1; -.
CC HSSP: P08631; IAD5.
CC InterPro: IPR00719; -.
CC InterPro: IPR001245; -.
CC InterPro: IPR003006; -.
CC Pfam: PF00047; I9; 3.
CC Pfam: PF00069; PKinase; 1.
CC PRINTS: PR00109; TYRKINASE.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.

```

```

DR SMART: SM00219; TYRKC; 1.
KW ATP-binding; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 830 AA; 92831 MW; 980E4CB0D4C1EA47 CRC64;

alignment_scores:
Quality: 452.50 Length: 385
Ratio: 1.926 Gaps: 10
Percent Similarity: 61.039 Percent Identity: 30.909

alignment_block:
us-09-598-042a-2 x 09UIN6
Align seg 1/1 to: 09UIN6 from: 1 to: 830

100 GTGGCGCGCGTGGCGCGCATATGCGGCTGACAGTCCGACGAGGAGGA 149
||||| .....TACGGCGCCAGGGCCGCCACACTC 863
51 ValAlaAlaProGlyGlnSerLeuGlnValArgCysLeuLys..... 65
150 CCGCGCGCGCGTACCATGTGGACCAAGATGGCGCACCATCCACAGCG 199
66 AspAlaAlaValIleSerTrpThrLysAspGly.....ValHisLeuG 80
200 GTGGAGCGCGTCCGCGTGTG...CCGACGGGCTGAGGTGAGAGG 246
80 LyrAsnAsnArgThrValLeuIleGlyGlyLysLeuIleLysGly 96
247 GTGAGCGGAGATGCGCGCGCTGTACGTGTGACAGCCGACCGACGCT 296
...TACGGCGCCAGGGCCGCCACACTC 863
97 AlaThrProArgAspSerGlyLeuTyrAlaCysThrAlaSerArgTyrVa 113
297 CCGCACGCTGAGCGTCACTACACCTGCTGCTGCATGACATTAGCC 346
113 AspSerGlnThrTrpTyrPheMetValAsnValThrAspAlaIleSerS 130
347 CAGGAGAGAGAGAGCTGGGGCGCGGACGCTCTGCGGGGTCAAGAGAG 396
130 ergLyspAsp.....GlnAspAspThrAspGlyAlaGlnLysp 142
397 CCGCGCGCGCAG.....CAGTGGCGACGACCGCGCGCTGCACAGGCC 437
143 PheValSerGlnAsnSerAsnAsnLysArgAlaProTyrTrpThrAsn 159
438 CTCACAGATGAGCGCGCGGTGATGCGACGCGCGCTGCTGCTGCTGCT 487
159 GlnLysMetGlnLysArgLeuHisAlaValProAlaAlaAsnThrValL 176
488 GGCTCAAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGATG 537
176 yspPheArgCysProAlaGlyGlyAsnProMetTrpThrMetArgTyrLeu 192
538 AAGGAGGACGACGCGCTTGACG.....CGCCGAGAGCGCGCTGAGGCC 578
193 LysAsnGlyLysGlnPheLysGlnGlnHisArgIleGlyGlyTyrLysVa 209
579 CAGGAAGAAGAAGTGGACACTGAGCGCTGAAGAAGCTGCGCGCGAGGACA 628
209 LArgSnglnHisTrpSerLeuIleMetGlnSerValValProSerAspL 226
629 GGGGCAATATACCTGCGCGCGCTGTGGAACCGCGCGCGCGCGCGCTCAAGCG 678
226 ysgLysnTyrThrCysValValGlnAsnGlyTyrGlySerIleAsnHis 242
679 ACCTCAAGGTGATGTGATCCAGCGAGCGCGCTGTCAAGCGCGCTGCTAC 728
243 ThrTyrHisLeuAspValValGlnArgSerProHisArgProIleLeuG 259
729 AGGCAAGCAGCCCGTGAACACAGCAGCGTGTGCGGGGACACAGCTGCT 778
259 nAlaGlyLeuProAlaAsnAlaSerThrValValGlyLysAspValGlnP 276
779 TGCATGCAAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 828
||| .....TACGGCGCCAGGGCCGCCACACTC 863

```

```

276 heValcysLysValTyrSerAspAlaGlnProHisIleGlnTrpIleLeu 292
829 CCGGTGAG.....TACGGCCCGGAGGCGCCGACCAATC 863
      |||||
293 HisValGlnLysAsnGlySerLysTyrGlyProAspLys..... 305
864 CACCATGATGTGGGGCGCCAGAAAGTTGTGTCGTCGACCGGTCAGC 913
306 .....LeuProTyrLeuLysV 311
914 TG.....TGTGCGCGCCGACGCGCTCTACCTCAATAAGCTGCTCATC 957
      |||
311 alleuLysHisSerGlyIleAsnSerSerAsnAlaGlnValLeuAlaLeu 327
958 ACCCGTCCCGGACGATCGGGGATGATCATCTGCTTGGCGCAA 1007
      |||||
328 PheAsnValThrGlnAlaAspAlaGlyLutryrLecysLysValSerAs 344
1008 CACCATGGGTACAGCTTCGCGACGCGCTTCCTACCGTGTGTCAGACC 1057
      |||||
344 nTyrIleGlyGlnAlaAsnGlnSerAlaTrpLeuThrValLeuProLysG 361
1058 CAACACCGCGAGGCGCACCTGTGGCTCTCTGTCCTGCGCCACTAGCCTG 1107
      |||||
361 lnglnAlaProGly.....ArgGlnLysGlnIleThrAlaSer 373
1108 CCGGCGCCCTGTCATCGGCATCCGACGCGCTCTTTCATCCGGG 1157
      |||||
374 ProAspTyrLeuGlnIleAlaIleTyrCysIleGlyValPheLeuIleAl 390
1158 CACCGCTGCTGCTGCTTGGCGCGCCAGAAAGCGCTGCACCCCGC 1207
      |||||
390 acysMetValValThrValIleLeuGlyArgMetLysAsnThrLysL 407
1208 CGCCT 1212
      |||
407 yspPro 408

```

seq\_name: sp\_rodent:Q00389

```

seq_documentation_block:
ID Q00389 PRELIMINARY; PRT; 820 AA.
AC Q00389;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE FIBROBLAST GROWTH FACTOR RECEPTOR 2 PRECURSOR (FGF-RECEPTOR).
GN FGFR2 OR BKR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/C; TISSUE=BRAIN;
RX MEDLINE=91270892; PubMed=1711190;
RA Raz V., Kejman Z., Avioli A., Neufeld G., Givol D., Yarden Y.;
RT "PCR-based identification of new receptors: molecular cloning of a
RT receptor for fibroblast growth factors.";
RL Oncogene 6:753-760(1991).
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
CC EMBL: X55441; CAA39083.1; -.
DR PIR: S17295; S17295.
DR HSSP: P06239; 3LCK.
DR MGD: MGI:95523; Fgfr2.
DR InterPro: IPR000719; -.
DR InterPro: IPR001245; -.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; Ig_3.
DR Pfam: PF00069; Kinase_1.
DR PRINTS: PRO0109; TYRKINASE.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.

```

```

DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR SMART: SM00219; TyrKc; 1.
KW ATP-binding; Signal; Transferase; Tyrosine-protein kinase.
FT SIGNAL 1 21
FT CHAIN 22 820 FIBROBLAST GROWTH FACTOR RECEPTOR 2.
SQ SEQUENCE 820 AA; 92060 MW; 5AC85E46952E1635 CRC64;

```

```

alignment_scores:
Quality: 447.50 Length: 378
Ratio: 1.946 Gaps: 9
Percent Similarity: 60.847 Percent Identity: 31.217

```

alignment\_block:

US-09-598-042a-2 x Q00389 ..

Align seg 1/1 to: Q00389 from: 1 to: 820

```

118 ACTATGGGCTGCAGTGCACAGTGAGAGGAGCCGCGCGCTGACCAT 167
      |||||
57 SerLeuGlnLeuGlnCysMetLeuLys.....AspAlaIleValIleSe 71
168 GTGACCAAGATGGCCGACCATCCACAGCGGCTGAGCCGCTCCGC 217
      |||||
71 rTrpThrLysAspGly.....ValHisLeuGlyProAsnAsnArgThrV 86
218 TGTGTG..CCGCAAGGGCTGAAGTGAAGTGAAGTGAAGCGGGAGATGCC 264
      |||||
86 alleuIleGlyArgTyrLeuGlnIleLysGlyValThrProAsnAspSer 102
265 GCGGTGACGTGTGCAAGGCGCACCAAGCGGCTTCGCGAGCTGACGTCA 314
      |||||
103 GlyLeuTyrIleAcysThrAlaAlaIleArgThrValAspSerGlnThrTrp 119
315 CTACACCTCTGCTGCTGATGATGATGATGATGATGATGATGATGATG 364
      |||||
119 rPheMetValAsnValThrAspAlaIleSerSerGlyAspAspGluAsp 136
365 GCGCCGACAGCTCTGCTGCGGCTCAAGAGAGCCCGCGACGACGATGG 414
      |||||
136 sPThrAspSerSerGluArgValSerGlnAsnArgSerAsnGlnArg... 151
415 GCAAGACCGGCTTCACACAGCCCTCCAGATGAGGCGCGGATGATGCC 464
      |||||
152 ...AlaProTyrTrpThrAsnThrGlnLysMetGlnLysArgLeuHisAl 167
465 ACGGCGCGGTGGTACCTCCGCTGCGGCTCAAGTGCCTGCGGCGGACCC 514
      |||||
167 aValProAlaAlaAsnThrValLysPheArgCysProAlaGlyGlyAsn 184
515 CTGCGCGCGACATCATCGTGTGATGAGAGAGCAGCAGCGCTGACG... 558
      |||||
184 roThrProThrMetArgTrpLeuLysAsnGlyLysGlnLysGlnGlu 200
559 ...CGCCAGAGCGCGCTGAGCCAGAGAGAGAGAGAGTGAAGTGAAGCT 605
      |||||
201 HisArgIleGlyGlyTyrLysValArgAsnGlnHisArgSerLeuIle 217
606 GAAGACCTGCGCGCGGAGAGAGAGCGGCAATACACTGCGCGGTGTCGA 655
      |||||
217 GluSerValValProSerAspLysGlyAsnIleThrCysLeuValGlu 234
656 ACGCGCGGCGCGCATCAAGCCATCAAGTGAAGTGAAGTGAAGTGAAGCG 705
      |||||
234 snGluTyrGlySerIleAsnHisThrLysIleAsnValValGluArg 250
706 ACCCGTTCAGACCGCTGCTCAAGAGCAGCAGCAGCAGCAGCAGCAGC 755
      |||||
251 SerProHisArgProIleLeuGlnAlaGlyLeuProAlaAsnAsnSerTh 267
756 GGACTTGGGGGAGCAGCTCTCCAGTGCAGTGCAGTGCAGGAGGAGCTGA 805
      |||||

```







```

||||| .....:|||||:
253 H1SHrYrH1SLeAspVal11LeGluArGSerSerH1sArGPro11Le 269
726 CAGAGGACGACGCCGAGACAGGAGGACTGCGGGGAGCCAGCT 775
269 uG1AlaG1LeuProAlaAsnThrThAlaValAlc1yLysAlaG 286
776 CCTTCAGTCAAGTGGCGGACGACGACGCGGTGATCCAGTGGCTG 825
286 LuPheValCysLysVal1yLrSerAspAlaGlnProH1s1LeArGTrpVal 302
826 AAGCGCGTGAAGTACGGCGCGGAGGCGGACCAACTCCACCATGATGT 875
303 ArGyTr11eGlu.....LysAsnG1SerArGpHeG1yVa 314
876 GCGCGGC...CAGAAGTTTGCTGCTGCCACGGGTACGTGCTGGC 922
314 lAspG1yLeuProTyTrPheLysVal1LeuLysH1sSerG1yLe..... 328
923 GCGCGGAGGCTCTACCTCAATAAGCTGTCATCACCGCGCGCCAG 972
329 .....AsnSerSerAlaG1uVal1LeuLysLeuTyTrAsnValThrGlu 343
973 GACGATGGGCGCATGTACATCTGCTGGCGCAACACCATGAGCTACAG 1022
344 AlAspAlaG1yGluTyTr11eCysThrValSerAsnTyTr11eG1yGluAl 360
1023 CTTCCGACGCGCTTCTCTACCGCTGCTGCCAGACCCAAACCGCCAGGCG 1072
360 AsnLysSerAlaTrPLeuMetValG1uArG1uLysAlaProValAsnP 377
1073 CA.....CCTGTGGCTCTCTGCTCTGCGCCATGAGCTGGCG 1110
377 rOlEuG1uAspAsnProVal.....Pro 384
1111 TGCGCGGTGTCATCGGCATCCAGCGCGCTGTCTCATCTGCGGAC 1160
385 TyTrYrMetG1u11eG1y11eTyTrSerThrG1y11ePhe11e11ePheCy 401
1161 CCTGCTCTCTGCGCTTGC.....CAGGCC 1186
401 sMeValVal...ValCysValValCysArGMe1ArG1nG1yAlaLysL 417
1187 AGAGAAGCCGTGACACCCCGCGCTGCCCTCCCTG..... 1224
417 ySLysLysAsnPhetHrG1yProProValH1sLysLeuThrLysArG11e 433
1225 CCTGGGACGCGCGCGGAGGAGCGCGGACGACGCGAC 1263
434 ProLeuH1sArG1nValThrValSerAlaAspSer 446

seq_name: sp_human:Q14308

seq_documentation_block:
ID Q14308 PRELIMINARY; PRT: 731 AA.
AC Q14308;
DT 01-NOV-1996 (Tremblrel. 01, Created).
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE NOVEL GROWTH FACTOR RECEPTOR (FRAGMENT).
DE FGFRL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RX MEDLINE=92147110; PubMed=1664411;
RX Thompson L.M., Plummer S., Schalling M., Altherr M.R., Gusella J.F.,
RA Hausman D.E., Wasmuth J.J.;
RT "A gene encoding a fibroblast growth factor receptor isolated from the
RT Huntington disease gene region of human chromosome 4.";
RL Genomics 11:1133-1142(1991).

```

```

CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL: M64347; AA58470.1; -.
DR HSP: P12931; 1FMK.
DR InterPro: IPR000719; -.
DR InterPro: IPR001245; -.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 1g; 3.
DR Pfam: PF00069; pkinase.1.
DR PRINTS: PR00109; TYRKINASE.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR SMART: SM00219; TyrKc; 1.
KW ATP-binding; Transferase; Tyrosine-protein kinase.
FT NON_TER 1
SQ SEQUENCE 731 AA; 80427 MW; F7B61CEAEAF6570D CRC64;

```

```

alignment_scores:
Quality: 438.00 Length: 413
Ratio: 1.752 Gaps: 13
Percent Similarity: 60.533 Percent Identity: 30.508

```

```

alignment_block:
US-09-598-042a-2 x Q14308 ..

```

```

Align seg 1/1 to: Q14308 from: 1 to: 731

```

```

175 AAGATGGCGCGCACCATTCACAGCGCTGAGCCGCTTCGCGCTGCTGCC 224
||||| .....:|||||:
1 LysAspG1yThnG1yLeuValProSer...GluArGVal1LeuValG1yPr 16
225 GCAGGGGCTGAAGTGAAGCAGTGGACGGAGGAGGATCGCGCTGTACG 274
16 cG1nArG1uG1nVal1LeuAsnAlaSerH1sG1uAspSerG1yAlaTyS 33
275 TGTGC....AAGCCACCAACAGGCTTCGGACGCTGACGCTCAACTAC 318
33 eCyArG1nArG1uArG1uThrGln.....ArgValLeuCyshHispe 46
319 ACCCGTGTGCTGTGATGACATTCACCGGAGGAGGAGGAGCGCTGGCGC 368
.....:|||||:
47 SerValArGValThrAspAlaProSerSerG1yAspAspG1yLysP1yG1 63
63 uAspG1uArG1uAspThnG1yValAspThrG1yAlaProTyTrP.... 78
419 GACCGCGCTTCACACAGCCCTCCAAAGATGAGCGCGCGGTGATCCAGG 468
||||| .....:|||||:
79 .....ThrArProG1uArG1uArG1uAspLysLysLeuLeuAlaVal 91
469 CCGCGGGTAGTCCGTCGCGGCTCAAGTGCCTGAGCGGAGGAGGAGCGCTG 518
||||| .....:|||||:
92 ProAlaAlaAsnThrValArGpHeArGysProAlaAlaG1yAsnProTh 108
519 GCCCGACATCAAGTGAAGTGAAGAGGAGCGCTTGACG.....C 559
||||| .....:|||||:
108 rProSer11eSerTrPLeuLysAsnG1yArG1uPheArG1yG1uH1sA 125
560 GCCCAGAGCGCGCTGACCCAG3AAGAAGAGTGAACACTGACCTGAAG 609
||||| .....:|||||:
125 rG11eG1yG1yL1eLysLeuArG1nG1nG1nTrPserLeuValMetG1u 141
610 AACCTGGCGCGGAGGAGCGG3CAATATACACTCCGCGGTGTGACAGCG 659
.....:|||||:
142 SerValValProSerAspArG1yAsnTyTrThrCysValAlc1uAsnHys 158
660 GCGGCGCGCATCAACAGCACCATTCACAGGTGATGTGATCCAGCGAGCC 709
||||| .....:|||||:
158 sPheG1ySer11eArG1nThrTyTrThrLeuAspVal1LeuG1uArGyserp 175

```

```

710 GTTCAAGCCCGTCTACAGAGCCACCGCCCTGACAGAGCGTGGAC 759
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
175 rohtsargprotleuvalaglyleuprolaasnlnthralaval 191
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
760 TTGGGGGGGACACGTCCTTCCAGTGCAGAGTGCAGAGCTGAAGCC 809
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
192 lenglyseraspvalglupheniscylsvaltyrseraspalaqlnpr 208
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
810 GGTATCCAGTGGCTGAGAGCGCTGAGTACAGCGCGCGGCGGCGAC 859
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
208 ohstlleglntrpleulyshtsvalglval.....asn 220
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
860 ACTCAACATCGATGTGGCGGCGCAAGATTGTG...GTGTCGCCACG 906
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
220 lysterlsvaglylproaspglythrprotyrvalthrvalleulyslth 236
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
907 GGTGAGCTGTGTGCGCGCGCGGCTCCTACCTCAATAGCTGTCTAT 956
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
237 AlaGly.....Alaasnthrtraspysgluleuglvalleuserle 251
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
957 CACCGCGCGCGCGAGAGCGAGTGCAGTACATCTGCTGGCGGCA 1006
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
251 uhtsaasnvalthrphelinspalaqlglutyrthrCysleuAlaGlyA 268
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
1007 ACACATGGGCTACAGCTCCGCGAGCGCTTCTCTCAACCTGTGCCAGAC 1056
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
268 snserlleglypheserhtshsersalatrpleuvalvalleuprola 284
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
1057 CCAAAACCGCGAGCGCGCTGTGGCTCTCTCTCGCGCCACTAGCCT 1106
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
285 gluglu.....gluleuvalgluAlaaspbluAlaGlyser.. 296
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
1107 GCCGTGGCGCGCTGATCGGATCCAGCGCGCGCTGTG..... 1146
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
297 .....ValTyrAlaGlylleuserltyrGlyValGlyPhePhe 310
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
1147 ..TTCATCTGGGCGACCTGCTCTGTGCTTGCAGGCGCGCAAGAG 1194
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
310 eupheilleuvalvalAlaAlaValthrValCysarGleuArGserPro 326
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
1195 CGGTGACCGCGCGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 1220
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
327 ProLysLysGlyLeuGlySerProthralHisLysIleSerArGpHeP 343
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
1221 CCGTCTGGGCGACCGCGCGCGGAGCGCGCGCGCGAGCGAGACA 1270
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
343 oleuLysarGlnvalSerleuGlnSerAsnAlaSerMetSerAsn 360
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
1271 AGGACCTTCCCTGCTGGCGCGCGCTGAGCGCTGCTGTGGGGCTG 1320
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
360 hrProleuValArgIleAlaArgLysSerSerGlyGlnGlyProthrl 376
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
1321 TGT.....GAGAGCATGGGTCTCGGCGAGCGCGCGCGAG 1353
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
377 AlaAsnValSerGlnLeuGlnLeuProAlaAspProLys 389
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
seq_name: sp_vertibrate:Q91286
seq_documentation_block:
ID Q91286 PRELIMINARY; PRT; 824 AA.
AC Q91286;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE FIBROBLAST GROWTH FACTOR RECEPTOR 2.
GN FGFR-2.
OS Pleurodeles waltl (Iberian ribbed newt).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidae; Salamandridae;
OC Pleurodeles.
OX NCBI_TaxID=8319;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=94299014; PubMed=8026621;
RA Shi D.L., Launay C., Fromentoux V., Felge J.J., Boucaut J.C.;
RT "Expression of fibroblast growth factor receptor-2 splice variants is
RT developmentally and tissue-specifically regulated in the amphibian
RT embryo."
RL Dev. Biol. 164:173-182(1994).
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL: X74332; CAA52379.1; -.
DR HSPSP; P08631; IAD5.
DR InterPro; IPR000719; -.
DR InterPro; IPR001245; -.
DR InterPro; IPR003006; -.
DR InterPro; IPR003598; -.
DR Pfam; PF00047; Ig; 3.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR SMART; SM00408; IgC2; 1.
KW ATP-binding; Transferase; Tyrosine-protein kinase.
SO SEQUENCE 824 AA; 92306 MW; 58B39786746D620B CRC64;

```

```

alignment_scores:
  Quality: 437.00      Length: 374
  Ratio: 1.884        Gaps: 13
  Percent Similarity: 62.032  Percent Identity: 30.214

```

alignment\_block:

us-09-598-042a-2 x Q91286

Align seg 1/1 to: Q91286 from: 1 to: 824

```

112 GCCGCACTATGCGGCTGCGATGCCAGTGGAGGGGACCGCGCGCT 161
    |||:||||| :|||:||||| :|||:||||| :|||:|||||
55 GlyLysSerValGlnLeuArGys...LeuLeuGlyAspAlaIleMetVa 70
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
162 GACCATGTGGACCAAGATGGCGCGCGCGCGCGCGCGCGCGCT 211
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
70 Lthr...TrrPthrLysAspGlyAlaGlnLeuValSerAsn...AsnArG 85
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
212 TCCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGAT 261
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
85 hrleuLeuIleGlyProtyrLeuGlnIleArgGlnAlaIleArProArG 101
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
262 GCCGCGGTATGCTGTGACAGCGCGCGCGCGCGCGCGCGCGCGCT 311
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
102 SerGlyLeuTyralCysGlyAlaThrArGProLeuGlnSer...GluTh 117
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
312 CAACATCACCCCTGCTGCTGATGACATTAGCCAGGAGGAGAGAGCC 361
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
117 rArgTyrrPheIleVal.....AsnIleThrAspGlyAsnSerSerG 131
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
362 TGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 411
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
131 LysAspAspGlnAspAspAsnAspGlySerGlnAspPheThrAsnAsp 147
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
412 TGGGCGAGA.....CGCGCTTACACAGCGCGCGCGCGCGCGCG 455
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
148 AsnHisLysAlaGlnAlaProtyrGlnThrAsnThrGlnLysLeuGln 164
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
456 GGTGATCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 505
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
164 SLeuHisAlaValProAlaAlaAsnThrValLysPheArGysProAlaG 181
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
506 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTG 555
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
181 LysGlnProLeuProSerMetArGTrPleuLysAsnGlnGlyGluPhe 197
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||
556 AGC.....CGCCAGAGCGCGCGCGCGCGCGCGCGCGAGAGAGAGT 596
    :|||:||||| :|||:||||| :|||:||||| :|||:|||||

```

```

198  Lys6IngluHnHsArg1IleGlyGlyPheLysValArgSerGlnHnIstrpse 214
597  ACTGAGCCTGAGAACCTGGCGCCGAGAGACAGCGCAATATACCTGCC 646
214  rleuileMetcluserValValProSerAspLysGlyAsnIlyThrCysI 231
647  GCGGTGGAGACCGGGGGGCGCCATCAAGCCACCTTCAAGGTGGAGTGTG 696
231  leMetGluAnsnGlyTrGlySerIleAsnHnIstrIlyHnIstrLysVal 247
697  ATCCAGCGGAGACCGCTTCCAAAGCCGCTGTACAGCAGCAGCAGCCGTA 746
248  ValGluArgSerTronHnIstrProIleGluGlnIleGlyLeuProAlaAs 264
747  CACGAGCGTGAAGCTTGGGGGAGACAGCAGCTTCCAGTCAAGGTGGCA 796
264  nThrThrThrLysValGlyGlyAspArgGluPheValLysLysValIlyrs 281
797  GCGACGTGAAGCCGCTGTATCCAGTGGCTGAAGCGCGTGGAGTACGGGCC 846
281  erAspAlaGlnProHnIstrIleGlnTrp..... 289
847  GAGGCGCCGACCACTCCACATGATGTGGCGCCGAGAGATTGTGTGT 896
290  ....ArgHnIstr.....IleGluLeuAsnIlySerLysIle..... 299
897  GCTGCGCCAGCGGTGACGTGTGGTGGCGCCGAGCGCTCTCACTCAATA 946
300  ....  GlyProAspGlyAsnProIlyLeuL 308
947  AGCGTC..... 954
308  ysaValLeuLysAlaAlaGlyLysAnThrThrValLysGluIleGluVal 324
955  ....ATCACCCGTGCCCGCCAGAGATGCGGAGTGTACATCTGCT 998
325  LeuThrValArgAnValSerPheGluAspAlaGlyLysIlyThrCysLe 341
999  TGGCGCCACACACATGAGGTACAGCTTCCGAGCGCTTCTCAACCGTGC 1048
341  uAlaGlyAsnSerIleGlyIleSerIlyHnIstrAlaTrpLeuThrValL 358
1049  TGGCAGACCCAAACCGCCAGCGCCACCTGTGGCTCTCTCTCTCGGCC 1098
358  euProAspGlyGluArgGlnLeuAsp.....SerSerSerSerGlyIlyTr 372
1099  ACTAGCCTGCGCGTGGCGCGTGCATCGGATCCAGCGCGCGCTGTCTT 1148
373  ThrGluIle.....AlaIlyTrp 378
1149  CATTCTGGGACACCTGTCTCTG 1170
378  rcysValAlGlyGlyPheLeuIle 385

seq_name: sp_human:Q12922
seq_documentation_block:
ID Q12922 PRELIMINARY; PRT; 366 AA.
AC Q12922;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE SOLUBLE KERATINOCYTE GROWTH FACTOR RECEPTOR PRECURSOR (EC 2.7.1.112).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BREAST, AND CORNEA;
RX MEDLINE=95170769; PubMed=7866434;
RA Wilson S.E., Weng J., Chwang E.L., Gollan L., Leitch A.M.,
RA Shay J.W.,

```

```

RT "Hepatocyte growth factor (HGF), keratinocyte growth factor (KGF), and
RT their receptors in human breast cells and tissues: alternative
RL receptors."
RL Cell. Mol. Biol. Res. 40:337-350(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-BREAST, AND CORNEA;
RX MEDLINE=92108030; PubMed=1309608;
RA Miki T., Bottaro D.P., Fleming T.P., Smith C.L., Burgess W.H.,
RA Chan A.M., Aaronson S.A.;
RT "Determination of ligand-binding specificity by alternative splicing:
RT two distinct growth factor receptors encoded by a single gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:246-250(1992).
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL: U11814; AAA68514.1;
DR InterPro: IPR003006;
DR InterPro: IPR003598;
DR Pfam: PF00047; Ig_3;
DR SMART: SM00408; IgC2; 1.
KW Receptor; Glycoprotein; Immunoglobulin domain; Signal; Keratin.
FT SIGNAL 1 21
FT CHAIN 22 366
FT FT
FT DOMAIN 55 114
FT DOMAIN 172 238
FT DOMAIN 271 349
FT DISULFID 62 107
FT DISULFID 179 231
FT DISULFID 278 340
FT CARBOHYD 83 83
FT CARBOHYD 123 123
FT CARBOHYD 128 128
FT CARBOHYD 241 241
FT CARBOHYD 241 241
FT CARBOHYD 265 265
FT CARBOHYD 297 297
FT CARBOHYD 318 318
FT CARBOHYD 331 331
SQ SEQUENCE 366 AA; 40614 MW; C02708836203465F CRC64;

alignment_scores:
Quality: 426.50 Length: 332
Ratio: 2.070 Gaps: 9
Percent Similarity: 62.048 Percent Identity: 32.530

alignment_block:
US-09-598-042a-2 x Q12922
Align seg 1/1 to: Q12922 from: 1 to: 366
100 GTGGCCGCGGTGGCGCCGATATGCGGTGCGTGCAGCGGAGAGGGA 149
||||| |||||
51 ValAlaAlaProGlyGlySerLeuGlnValArgCysLeuLeuLys..... 65
150 CCGCGCGCGCTGACATGTGACCAAGATGCGCGACCATCCACAGCG 199
||| |||||
66 AspAlaAlaValIleSerTrpThrLysAspGly.....ValHnIstr 80
200 GGTGAGCGCGCTTCCGCGTGTG...CCGAGGCGGTGAGTGAAGAG 246
||| |||||
80 LysProAsnAsnArgThrValLeuIleGlyGlyTrpLeuGlnIleLysGly 96
247 GTGAGCGCGGAGATGCGCGCTGTACGTGTGCAAGCGCACCAAGCGGTT 296
||| |||||
97 AlaThrProArgAspSerGlyLeuThrAlaCysThrAlaSerArgThrVal 113
113 LaspSerLysThrTrpTrpPheMetValAsnValThrAspAlaIleSer 130
347 CAGGAGAGAGAGCGCTGGCGCGCGACAGCTCTCTGGGGGTCAAGAGAC 396
||| |||||

```

```

130 erc1yaspasp.....GluAspAspThrAspGlyAlaGluAsp 142
397 CCGCCAGCCAG.....CAGTGGCAGACCGCGCTTACACAGCC 437
143 phevalsergluasnserasnlnlyslargalaProtylTrpThrAsnTh 159
438 CTCGAGATGAGCGCGGATGATCCACGCGCGGAGTACGCGTGC 487
159 rglulysmetglulysargleuHisAlaValProAlaAlaAsnThVal 176
488 GGCTCAAGTGGCTGGCAGCGCGCGCTGGCGGACATCACGTGATG 537
176 yshargysProAlaGlyGlyAsnProMetProThrMetArgTrpLeu 192
538 AAGGAGCAGCGCGCTTGACG.....CGCCAGAGCGCGCGCTGAGCC 578
193 LysAsnGlyLysGluPheLysGlnGluHisArgIleGlyLysTrpVal 209
579 CAGGAGAGAGAGTGGACACTGAGCTGAGAGACTGGCGCGGAGAGCA 628
209 LArgAsnGlnHisTrpSerIleuIleMetGluSerValValProSerAsp 226
629 GCGGCAAAATACACCTGCGCGGTGTCGACCGCGCGCGCATCAAGCC 678
226 ysglYasnlyTrpThrCysValValGluAsnGluTrpGlySerIleAsnHis 242
679 ACCTACAGGTGATGTGATCCACGCGCGCTTCCAGCGCGCTGCTCAC 728
243 ThrTyrlHisLeuAspValValGluArgSerProHisArgProIleuG 259
729 AGGACGACGACCGCGCTGACAGACGAGTGGACTTGGGCGGAGCAGCT 778
259 nAlaGlyLeuProAlaAsnAlaSerThrValValGlyGlyAspValGlu 276
779 TCAGTGCAGAGTGGCGCGACGAGCTGAGCGCGGTGATCCAGTGGCTG 828
276 heValCysLysValIlySerAspAlaGlnProHisIleGlnTrpIleLys 292
829 CGCGTGGAG.....TAGGCGCGGAGGCGCGCGCGCACAACTC 863
293 HisValGluLysAsnGlySerLysTrpGlyProAspGly..... 305
864 CACCATGATGTGGCGCGCGCAGAGTTTGTGTCTGCTCCCGACGGGTG 913
306 .....LeuProTyrlLeuLysV 311
914 TG.....TGTGCGCGCGCGCGCGCTTACCTCAATAAGCTGCTATC 957
311 alleuLysHisSerGlyIleAsnSerSerAsnAlaGluValIleuAlaLeu 327
958 ACCGTCGCGCGCGCGAGAGATGGCGGATGATGATGCTGCGCTGGCGCA 1007
328 PheAsnValThrGluAlaAspAlaGlyGluTrpIleCysLysValSerAs 344
1008 CACCATGAGGCTACAGCTTCCGACGCGCTTCCACCGCTGCGCA 1053
344 nTyrlIleGlyAlaAsnGlnSerIleAlaTrpLeuThrValLeuPro 359
seq_name: sp_r0dent.Q60818

```

```

seq_documentation_block:
ID 060818 PRELIMINARY; PRT: 822 AA.
AC 060818;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 16, Last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE FIBROBLAST GROWTH FACTOR RECEPTOR-1, LONG ISOFORM PRECURSOR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBL_TaxID-10090;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=SWISS WEBSTER; TISSUE=HEART;
RX MEDLINE=95205422; PubMed=7897669;
RA Jin Y., Pasumamathi K.B.S., Bock M.E., Lytras A., Kardam E.,
RA Cattini P.A.;
RT "Cloning and expression of fibroblast growth factor receptor-1
RT isoforms in the mouse heart: evidence for isoform switching during
RT heart development.";
RL J. Mol. Cell. Cardiol. 26:1449-1459(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS WEBSTER; TISSUE=HEART;
RA Jin Y.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases
CC -1 SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL: U2324; AAC52182.1; -.
DR HSPB; P06239; 31CK
DR InterPro; IPR000719; -.
DR InterPro; IPR001245; -.
DR InterPro; IPR003006; -.
DR Pfam; PF00047; 1g; 3.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM_1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR_1.
DR SMART; SM00219; TykC; 1.
KW ATP-binding; Transferase; Tyrosine-protein kinase.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 822 FIBROBLAST GROWTH FACTOR RECEPTOR-1, LONG
FT FT ISOFORM.
SQ SEQUENCE 822 AA; 91938 MW; A6B6ABBAED450F5 CRC64;

```

```

alignment_scores:
Quality: 425.50 Length: 443
Ratio: 1.588 Gaps: 15
Percent Similarity: 60.497 Percent Identity: 27.991

```

```

alignment_block:
US-09-598-042a-2 x Q60818 ..

```

```

Align seg 1/1 to: Q60818 from: 1 to: 822
112 GCGCGCATATGCGCGCTGACATGCGCGAGTGAGGAGGCGCGCGCT 161
48 GlyAspleuLeuGlnLeuArgCysArgLeuArgAspValGlnSerI 64
162 GACCATGTGACCAAGATGCG...CGCACATCCACACGCGCTGAGCC 208
64 easn...TrpLeuArgAspGlyValGlnLeuValGluSer.....Asn 78
209 GCTTCCGCGGTGCGCGCGAGGCGTGAAGTGAAGAGTGAAGCGGAG 258
78 rghrArgIleThrGlyGluGlnValGluValArgAspSerIleProAla 94
259 GATCGCGCGCTGTACGTGTGCAAGCGCACACGCGCTTGGCGCTGAG 308
95 AspSerCysLeuTyrlAlaCysValThrSerSerProSerGlySer...As 110
309 CGTCACTACACCTGCTGCTGCTGCTGATGATGATGATGATGATGATG 358
110 pThrThrTyrlPheSerValAsnValSerAspAlaLeuProSerSerGlu 127
359 GCGTGGGCGCGCGAGCTCTGTGGGAGTCAAGAGAGACCGCGCC..... 402
127 spAspAspAspAspAspAspSerSerSerCysGluGluLysGluThrAsp 143
403 .....AGCAGAGTGGGCGACAGCGCGCTTACACAGCGCTTCCAA 443
144 ThrLysProAsnArgArgProValAlaIleProTyrlTrpThrSerProG 160
444 GATGAGCGCGCGGTGATGCGACGCGCGCGTGGAGTCTGCTGCGCTCA 493

```



Align seg 1/1 to: 002063 from: 1 to: 662

```

112 GCGCCGACTATGCGGCTGCATGCTCCAGTGAAGGGAGCCGCCGCCGT 161
    ||| :|||:||||| ||| :|||:||||| ||| :|||:||||| ||| :|||:|||||
48 GAApLeuLeuGlnLeuArgCysArgLeuArgAspValGlnSer11 64
162 GACCATGTGGACCAAGATGGCGCCGACCATCCACAGCGCTGAGCCGCT 211
    ||| :|||:||||| ||| :|||:||||| ||| :|||:||||| ||| :|||:|||||
64 easn...TTPLeuArgAspGlyValGlnLeuAlaGluSer...AspArgT 79
212 TCCGCGTGTGCGCGGAGGCTGAAGCTGAAGAGCTGAGCGGAGGAT 261
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
79 hArgylethrglygluValGlnValGlnAspSerValProAlaAsp 95
262 GCGCGGCTGACGCTGCAAGGCCCAACGCGCTTGCAGCGCTGAGCT 311
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
96 SerGlyLeuValAlaCysValThrSerProSerGlySer...AspTh 111
312 CAACATACACCTGCTGCTGATGACATTAACCCAGGAGAGAGAGCC 361
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
111 rThrTYrPheSerValAsnValSerAspAlaLeuProSerSerGluAspA 128
362 TGGGCGCCGACGCTCTGCGGGGTCAAGAGAGCCCGCC 402
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
128 spAspAspAspAspAspSerSerSerGluGluGluGluThrAspAsnThr 144
403 .....AGCCAGCATGGGCGACAGCGCGCTTGCAGCGCTTGCAGAT 446
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
145 LysProAsnArgMetProValAlaProTYrTYrThrSerProGluLysMe 161
447 GAGCGCGGCTGATGCGACGCGCGGTGATCTCGCGCTGAGTCAAGT 496
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
161 GAluLysLysLeuHisAlaValProAlaAlaLysThrValLysPheLysC 178
497 GCCTGGCGAGCGGCAACCTCGCGCGGACATCACTGATGAAGAGAC 546
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
178 YsProSerSerGlyThrProAsnProThrLeuArgTYrLeuLysAsnGly 194
547 CAGGCTTGACGGCCGACAG.....GCCGCTGAGCCGCGGAA 584
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
195 LysGluLysHe...LysProAsnHisArgGlyGlyLysTYrLysValArgTY 210
585 GAGAAATGACACTGACCTGAAGAAGCTCGCGCGGAGAGAGAGCGCA 634
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
210 rAlaThrTYrSerLeuLeuMetAspSerValValProSerAspLysGlyA 227
635 AATACACCTGCGCGCTGCAACCGCGGCGGCGCATCAAGCCGCTAC 684
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
227 snTYrThrCysLysValGlnLysAsnGluTYrGlySerThrLeuAsnHisThrTY 243
685 AAGGTGATGTGATCCAGCGGAGCCGCTCCAGCGCGTGTCCAGAGCAC 734
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
244 GlnLeuAspValValGlnArgSerProHisArgProThrLeuGlnAlaG 260
735 GCACCCGCTGAACAGAGGCTGACTTCGGGGGAGACACGCTCTTCAGT 784
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
260 yLeuProAlaAsnLysThrValAlaLeuGlySerAsnValGlnPheMetC 277
785 GCAAGGTGGCGAGCGACTGAAAGCGGTGATCCAGTGGCTGAAGCGGTG 834
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
277 yLysValTYrSerAspProGlnProHisLysGlnTYrPheLysHisLys 293
835 GAGTACGGCGCGGAGGCGCGCACAACTCCACATCATGATGTTGGGCGCA 884
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
294 Glu.....ValAsnGlySer 298
885 GAAGCTTT.....GTGTCGCTGCCAGCGGTG 910
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
298 rLysIleGlyProAspAsnLeuProTYrAlaGlnIleLeuLysThrAlaG 315
911 ACGTGTGTCGCGCGCGAGCGCTCTTACCTCATTAAGCTCTCATACACC 960
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||

```

```

315 LysVal.....AsnThrThrAspLysGlnMetGlnValLeuHisLeuArg 329
961 CGTGCCCGCCAGGAGACATGGCGGATGTACATCTGCTGGCGCCACAC 1010
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
330 AsnValSerPheGlnAspAlaGlyGluTYrThrCysLeuAlaGlyAsnSe 346
1011 CATGGCTTACACCTTCCGACGCGCTTCCGACGCTGCTGCCA.....G 1054
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
346 rLecGlyLeuSerHisHisSerAlaThrPheThrValLeuGlnAlaLeuG 363
1055 ACCCAAAACCGCCAGGCGCCACTGTGGCTCTCTCTCGCGCCACTAGC 1104
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
363 LysGluArgPro.....AlaValMetThrSer 371
1105 CTGCGGTGGCGCGGTGTCATCGCATCCAGCGCGGCTGCTTCATC.. 1152
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
372 ProLeuTYrLeuGlnIleLeuLeuTYrCysThrGlyAlaPheLeuIleSe 388
1153 .....CTGGGACCCCTGCTCTGCTGCTTGCAGGCGCCGAAAGAGC 1195
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
388 rCysMetValGlySerValIleValTYrLysMetLysSerGlyThrLysL 405
1196 CG.....TGCACCCCGCGCGCTCCCGCTCCCGCG 1224
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
405 ySerAspPheHisSerGlnMetAlaValHisLysLeuAlaLysSerIle 421
1225 CTTGGGACCGCGCGCGCGGAGCGCGCCGACCGCAGCGGAGAGAGA 1274
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
422 ProLeuArgArgGlnValThrValSerAlaAspSerSerAlaSerMetAs 438
1275 C.....CTTCCTCTGCTGGCGCGCGCTCAGCGCTCAGCGTGGCG 1306
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
438 nSerGlyValLeuLeuValArgProSerArgLeuSerSerSerGlyThrP 455
1307 CTGGTGTGGGGCTGTGTGAGGAGCATGCTCGCGGACCGCCGAG 1353
    |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
455 rMetLeuAlaGlyValSerGlyTYrGluLeuProGlnAspProArg 470
seq_name: sp_Invertebrate:Q26614
seq_documentation_block:
ID Q26614 PRELIMINARY; PRT: 972 AA.
AC Q26614;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE FIBROBLAST GROWTH FACTOR RECEPTOR PRECURSOR.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidae; Euechinoidae; Echinacea; Echinoida; Strongylocentrotidae;
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RA Mccoon P.E., Angerer L.M., Angerer R.C.;
RL Submitted (NOV-1994) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
CC EMBL: U17164; AAC47258.1; -.
DR HSSP: P08631; IAD5.
DR InterPro: IPR000719; -.
DR InterPro: IPR001245; -.
DR InterPro: IPR001777; -.
DR InterPro: IPR003006; -.
DR Pfam: PF00041; fn3; 1.
DR Pfam: PF00047; ig; 3.
DR Pfam: PF00069; pktnase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR SMART: SM00219; Tyrc; 1.
KW ATP-binding; Signal; Transferase; Tyrosine-protein kinase.

```







alignment\_block:  
US-09-598-042A-2 x 09QW78

Align seg 1/1 to: 09QW78 from: 1 to: 376

```

112 GCGCGCATATGCGCTGCTGACCTGACCGGAGGAGGAGCCGCGCGCT 161
    ||| :|||:||||| :||| :||| :||| :||| :||| :||| :|||
48 GLYSPLEULEUGINLEUARGCYSEARGLEUARGSPASPAVALGINSEI 64
162 GACCATGTGGACCAAGATGGC...CGCACATCCACAGCGGCTGGAGCC 208
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
64 easn...TrrleuArgspGlyValGlnLeuValGlnSer.....AsnA 78
209 GCTTCGCGGTCTGCCGAGGGGCTGAAAGTGAAGCAACAGCTGAGCGGAG 258
    ||| :|||:||||| :||| :||| :||| :||| :||| :||| :|||
78 rgnrArGlieThrGlyGlnValGlnValArgspSerIleProAla 94
259 GATCCGCGCTGTAGTGTGCAAGGCGCACAAAGCGCTCGGACGCTGAG 308
    |||:|||||:||||| :||| :||| :||| :||| :||| :||| :|||
95 AsperGlyLeuTyAlaCysValThrserserProserGlySer..As 110
309 CGTCAACTACACCTCTGCTGTGATGATGATAGCCGAGGAGAGAGA 358
    :|||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
110 pThrThrTyPheSerValAsnValSerAspAlaLeuProserSerGluA 127
359 GCGTGGGGCGCGACAGCTCTCTGGGGGTCAAGAGACCCCGCC... 402
    ||| :|||:||||| :||| :||| :||| :||| :||| :||| :|||
127 spAspAspAspAspAspAspSerSerSerGlnGlnGlySglThrAspAsn 143
403 .....AGCCAGAGTGGGACGCGCGCTTACACAGCCCTCCAA 443
144 ThrIysProAsnArgArgProValAlaProTyTrpThrIserProGlu 160
444 GATGAGGCGCGGTGATGCGACGCGCGGTGAGTCTCGCGGCTCA 493
    |||:|||||:||||| :||| :||| :||| :||| :||| :||| :|||
160 smecGlnLysIleuHisAlaValAlaProAlaLysThrValLysPheL 177
494 AGTCGCTGGCGAGCGGCGACCTCGCGCGCGACATCAGCTGATGAAGAC 543
    |||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
177 yscysProserSerGlyThrProAsnProThrLeuArgTrpLeuLysAsn 193
544 GACGAGGCTTGACGCGCGCCAGAG.....GCCGCTGAGCCGAG 581
194 GlyLysGluPhe...LysProAspHisArgIleGlyGlyTyLysValAr 209
582 GAAGAGAGAGTGAACATGACCTGAAGAACCTGCGCGGAGAGACAGCG 631
    |||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
209 gTyAlaThrTrpSerIleIleMetLaspSerValValProserAspLysG 226
632 GCAATATACACTGCCCGCTGTCGAACCGCGCGCGCGCATACAGCCAC 681
    |||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
226 lYasnTyTrhCysIleValGlnAsnGlnTyTrhGlySerIleAsnHisP 242
682 TACAAGGTGATGTGATCCAGCGCGCGCTCCAAAGCGCGGCTCACAGG 731
    |||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
243 TyrlneUasPvalValGlnArgSerProHisArgProHisProSerAl 259
732 CACGACCCCGCTGAACAGACGTGACTGCGGGGAGCAAGCTCTTCC 781
    |||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
259 agLyleuProAlaAsnLysThrValAlaLeuGlySerAsnValGlnPheM 276
782 AGTCAAGGTGCGCAGCGAGCTGAAGCGCGGTGATCAGTGGCTGAAGCG 831
    |||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
276 etCysLysValTyTrhSerAspProGlnProHisIleGlnTrpLeuLys 292
832 GTGAGGTACGCGCGCGAGCGCGCGCAACATCCACCATGATGTGGGGG 881
    |||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
293 lIeclVal.....AsnGlySerLysIleGlyLysProAspAs 304
882 CCAGAGATTGTG...GTGTCGCCACGAGGTGAGCTGTGTCGCGCGCGG 928
    |||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
304 nLeuProTyValGlnIleLeuLysThrAlaGlyVal.....AsnThrT 319

```

```

929 AGCGCTCCATCATATAGCTGTCTATCATCCCGCGCGGAGGAGCAT 978
    :||| :|||:||||| :||| :||| :||| :||| :||| :||| :|||
319 hrAspLysGlnMetGlnValLeuHisIleLeuArgAsnValSerPheGln 335
979 GCGGCGATGATACATCTGCTGGCGCGCAACACCATGAGCTACAGCTTCC 1028
    |||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
336 AlaGlyGlnTyTrhCysLeuAlaGlnLysSerIleGlyLeuSerHisH 352
1029 CAGCGCTTCTCTGACCGCTGCTG 1050
    :|||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
352 sSerAlaTrpLeuThrValLeu 359
seq_name: sp_rodent:Q63242

```

```

seq_documentation_block:
ID Q63242 PRELIMINARY; PRT; 353 AA.
AC Q63242;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HEPARIN-BINDING FIBROBLAST GROWTH FACTOR RECEPTOR 2 (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PROSTATE;
RX MEDLINE=93326167; PubMed=8333865;
RA Van G., McBride G., McKeelhan W.L.;
RT "Exon skipping causes alteration of the COOH-terminus and deletion of
RT the phospholipase C gamma 1 interaction site in the FGF receptor 2
RT kinase in normal prostate epithelial cells.";
RL Biochem. Biophys. Res. Commun. 194:512-518(1993).
CC -!- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL: L19112; AAA02694.1; -.
DR InterPro: IPR003006; -.
DR InterPro: IPR003598; -.
DR Pfam: PF00047; Ig; 2.
DR SMART: SM00408; IGG2; 1.
KW Alternative splicing.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 353
SQ SEQUENCE 353 AA; 39140 MW; 5F54FC49D92B7FE CRC64;

```

alignment\_scores:  
Quality: 411.00 Length: 353  
Ratio: 1.995 Gaps: 8  
Percent Similarity: 58.357 Percent Identity: 31.728

alignment\_block:

US-09-598-042A-2 x Q63242

Align seg 1/1 to: Q63242 from: 1 to: 353

```

202 TGGAGCCGCTTCCGCTGCTGCCGCGAGGGGCTGAAGTGAAGCAGTGGA 251
    |||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
28 TrpGlyArgPheIleCysLeu..... 34
252 GCGGAGAGATGCCGGCGGTGATCGTGAAGGCCACCAAGCGCTTGGCG 300
    ||| :||| :||| :||| :||| :||| :||| :||| :|||
35 .....ValLeuValThrMetAlaThrLeuSerLeuAlaA 46
301 .....AGCTGAGCGCTACACCTGATGATGATGATGATGATGATGATG 339
    |||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
46 rGProSerPheSerLeuValGlnAspThrThrLeuGlnProGlnAspAla 62
340 ATTACCCAGGAGAGAGAGAGCTGGGGCGCGACAGCTCTGTGGGGGTCA 389
    |||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
63 lIeSerSerGlyAspAspGlnAspAspThrAspSerSerGlnAspPheVa 79

```

390 AGAGACCCCGCCAGCCAGTGGCGCAGACCGCGCTTACACAGCCCT 439  
: : : : : ||| : : : ||| : : : : : : : : : : : : : : : :  
79 lSerGluAsnArgSerAsnGlnArgAla...ProTyrTrpIleAsnThrG 95  
440 CCAAGATGAGCGCCCGGTGATCGCACGGCCGTGGTACTGCTCGG 489  
: : : : : ||| : : : ||| : : : : : : : : : : : : : : : :  
95 lIuysMetGluIuysArgLeuHisAlaValProIleAlaIleAsnThrValLys 111  
490 CTCAGTGCCTGGCCAGCGGGCCACTCGGCCGACATCACGTGATGAA 539  
: : : : : ||| : : : ||| : : : : : : : : : : : : : : : :  
112 PheArgCysProIleAlaGlyIleAsnProIleProIleMetArgTrpLeuIu 128  
540 GGACGACCCAGCCCTTGACG...GGCCAGAGCCCGCTGAGGCCA 580  
| : : : : : : : : : : : ||| : : : : : : : : : : : : : : :  
128 sAsnGlyLysGluPheLysGlnGluHisArgIleGlyIleTyrLysValA 145  
581 GGAAGAGAGATGGACACTGAGCCTGAAGACCTGGCGCGGAGACAGC 630  
| : : : : : ||| : : : ||| : : : : : : : : : : : : : : : :  
145 rGAsnGlnHisTrpSerLeuIleMetGluSerValValProSerAspLys 161  
631 GGCAAATACACTGCGCGGTGTCGACACCGCGCGCGCCATCAACGCCAC 680  
| : : : : : ||| : : : ||| : : : : : : : : : : : : : : : :  
162 GluAsnTyrThrCysLeuValGluAsnGluTyrGlySerIleAsnHisTh 178  
681 CTACAGGTGATGTGATCCAGCGACCCGCTCCAAACCCGCTGCTCACAG 730  
| : : : : : ||| : : : ||| : : : : : : : : : : : : : : : :  
178 rTyHisLeuAspValValGluArgSerProHisArgProIleLeuGlnA 195  
731 GCACGACCCCGGTGAACACAGCGGTGACTCGGGGGACACAGCTCTTC 780  
: : : : : ||| : : : ||| : : : : : : : : : : : : : : : :  
195 lAgIleuProAlaAsnAlaSerThrValValGlyGluAspValGluPhe 211  
781 CAGTGCAGAGTGGCGCAGCGTGAACCGGTGATCCAGTGGCTGAAGCG 830  
| : : : : : ||| : : : ||| : : : : : : : : : : : : : : : :  
212 ValCysLysValTyrSerAspAlaGlnProHisIleGlnTrpIleLysH 228  
831 CGTGGAG...TAGGGCGCCGAGGGCGGCCACAACTCCA 865  
: : : : : ||| : : : ||| : : : : : : : : : : : : : : : :  
228 sValGluLysAsnGlySerLysTyrGlyProAspGly..... 240  
866 CCATCGATGTGGCGCCACAGATTGTGCTGCCACAGCGGTGACGTG 915  
241 .....LeuProTyrLeuLysVal 246  
916 TGGTGGGGCGCCGACGGCTCTACCTCAATAG.....CTGCT 953  
: : : : : ||| : : : ||| : : : : : : : : : : : : : : : :  
247 LeuLysAlaIleAlaGlyValAsnThrThrAspLysGluIleGluValLeuTy 263  
954 CATCACCCGTGCGCCGAGAGATGGCGGCGATGTACATGTGCTTGGCG 1003  
| : : : : : ||| : : : ||| : : : : : : : : : : : : : : : :  
263 rIleArgAsnValThrPheGluAspAlaGlyGluTyrThrCysLeuAlaG 280  
1004 CCAACACCATGGGCTACAGCTTCCGACGCGCTTCCTACCGCTGTGCCA 1053  
: : : : : ||| : : : ||| : : : : : : : : : : : : : : : :  
280 lYAsnSerIleGlyIleSerPheHisSerAlaTrpLeuThrValLeuPro 296  
1054 GACCCAAAGCGCCAGCGCCACACTGTGGCTCTGCTCCGCGCCACTAG 1103  
| : : : : : ||| : : : ||| : : : : : : : : : : : : : : : :  
297 AlaPro.....ValArgGluLysGluIleThrAl 306  
1104 CTGCGCTGGCCCGTGTGATCGGATCCAGCGCGCGCTGTTCATCC 1153  
: : : : : ||| : : : ||| : : : : : : : : : : : : : : : :  
306 aserProAspTyrIleuGluIleAlaIleTyrSerIleGlyValPheLeuI 323  
1154 TGGGACCCCTGCTCTGCTTGGCTTCCAGCGCCACAGAGAACCGGTGCAC 1203  
: : : : : ||| : : : ||| : : : : : : : : : : : : : : : :  
323 leAlaCysMetValValThrValIlePheCysArgMetLysThrThrThr 339  
1204 CCGCGGCCT 1212  
| : : : : : ||| : : : ||| : : : : : : : : : : : : : : : :  
340 LysLysPro 342



[illegible]

```

US-09-383-586-24
; Sequence 24, Application US/09383586
; Patent No. 6242419
; GENERAL INFORMATION:
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Mathew
; APPLICANT: Abernethy, Nevlin
; APPLICANT: Onrust, Rene
; APPLICANT: Kumble, Anand
; APPLICANT: Murison, Greg
; TITLE OF INVENTION: Compounds isolated from stromal cells
; TITLE OF INVENTION: and methods for their use
; FILE REFERENCE: 11000.1037c1
; CURRENT APPLICATION NUMBER: US/09/383,586
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 1004
; TYPE: DNA
; ORGANISM: Human
US-09-383-586-24

```

[illegible]









```

Db 287 CCGGCTCTATGCTTGCATACAGCAGCCCTCGGGCAGTGCACACACTACTTCTCG 346
QY 323 tctgtctgtatgacatagaccaggaagagagccttgaggcccgacgtctctcg 382
Db 347 TCATATTTTCAGATGCTCTCCCTCTCGAGATGATGATGATGATGATGATGATGAT 406
QY 383 ggggtcaagagagagagagagagagagagagagagagagagagagagagagagag 430
Db 407 CAGAGAGAAAGAAAGATTAACACCAACCAACCAACCAACCAACCAACCAACCAACCA 466
QY 431 caagccctcaagatgagagagagagagagagagagagagagagagagagagagag 490
Db 467 CATCCCAAGAAAGATGAGAAAGAAATGCAATGCAATGCAATGCAATGCAATGCAAT 526
QY 491 tcaagtgctgtgagagagagagagagagagagagagagagagagagagagagag 550
Db 527 TCATAATGCCCTTCCAGTGGAGACCCCAACCCCAACCTGCGTGGTGGAAATGGCAAG 586
QY 551 ccttgagcgc-----gccagagagcgcgtgagcccaagagaagatgagacactga 601
Db 587 AATTCAACCTGACACACAGATTTGGAGGCTACAGGCTCCGTTATGCCACTGGAGCATCA 646
QY 602 gacctgaagaacctgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 661
Db 647 TAATGACTCTGTGTGCTGCTGACAGGCAAGGCAACTACACTGATTTGGAGAAATGAGT 706
QY 662 cggcgccatcaagcgcacactaaagtgatgagatgagatgagatgagatgagatgag 721
Db 707 ACGGCAGCATCAACACACATATACACTGATGATGATGATGATGATGATGATGATGAT 766
QY 722 tgcctcaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 781
Db 767 TCTTGAAGAGAGGTTGCCCGCAACAAACAGTGGCCCTGGGTACACAGTGGATTTCA 826
QY 782 agtgcagagtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 837
Db 827 TGTGTAAGGTGATGAGTACCCGACCGACGACATCCAGTGGCTAAAGCAATCGAG 882

```

RESULT 8  
 US-08-371-001-14  
 ; Sequence 14, Application US/08371001  
 ; Patent No. 5783683  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Morrison Ph.D., Richard  
 ; TITLE OF INVENTION: Methods and Composition for Treating  
 ; TITLE OF INVENTION: Tumor Cells  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Gray Cary Ware & Freidenrich  
 ; STREET: 401 "B" Street, Suite 1700  
 ; CITY: San Diego  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 92101  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/371,001  
 ; FILING DATE: January 10, 1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Brozman, Harris F.  
 ; REGISTRATION NUMBER: 35,461  
 ; REFERENCE/DOCKET NUMBER: P00095U0  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (619) 699-3630  
 ; TELEFAX: (619) 236-1048  
 ; INFORMATION FOR SEQ ID NO: 14:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 2733 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 210..467
; OTHER INFORMATION: REF1 Alpha Exon
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 57
; OTHER INFORMATION: "IDENTITY OF
; OTHER INFORMATION: NUCLEOTIDE PROVISIONAL"
; NAME/KEY: misc.feature
; LOCATION: 117
; OTHER INFORMATION: "IDENTITY OF
; OTHER INFORMATION: NUCLEOTIDE PROVISIONAL"
US-08-371-001-14

```

```

Query Match 3.5%; Score 110.8; DB 1; Length 2733;
Best Local Similarity 50.8%; Pred. No. 1.6e-16;
Matches 333; Conservative 0; Mismatches 302; Indels 21; Gaps 2;

```

```

QY 203 ggaagccgtctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 262
Db 344 GCAACCGCACCCGACATACACAGGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 403
QY 263 cggcgctgacgtgagcagagcagcagcagcagcagcagcagcagcagcagcagcagc 322
Db 404 CCGGCTCTATGCTTGCATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 463
QY 323 tctgtctgtatgacatagaccaggaagagagagccttgaggcccgacgtctctcg 382
Db 464 TCATATGTTTCAGATGCTCTCCCTCTCTCGAGAGATGATGATGATGATGATGATGATGAT 523
QY 383 ggggtcaagagagagagagagagagagagagagagagagagagagagagagagagag 430
Db 524 CAGAGAGAAAGAAAGATTAACACCAACCAACCAACCAACCAACCAACCAACCAACCA 583
QY 431 caagccctcaagatgagagagagagagagagagagagagagagagagagagagagag 490
Db 584 CATCCCAAGAAAGATGAGAAAGAAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 643
QY 491 tcaagtgctgtgagagagagagagagagagagagagagagagagagagagagagag 550
Db 644 TCATAATGCCCTTCCAGTGGAGACCCCAACCCCAACCTGCGTGGTGGAAATGGCAAG 703
QY 551 ccttgagcgc-----gccagagagcgcgtgagcccaagagaagatgagacactga 601
Db 704 AATTCAACCTGACACACAGATTTGGAGGCTACAGGCTCCGTTATGCCACTGGAGCATCA 763
QY 602 gacctgaagaacctgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 661
Db 764 TAATGACTCTGTGTGCTGCTGACAGGCAAGGCAACTACACTGATTTGGAGAAATGAGT 823
QY 662 cggcgccatcaagcgcacactaaagtgatgagatgagatgagatgagatgagatgag 721
Db 824 ACGGCAGCATCAACACACATATACACTGATGATGATGATGATGATGATGATGATGATGAT 883
QY 722 tgcctcaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 781
Db 884 TCTTGAAGAGAGGTTGCCCGCAACAAACAGTGGCCCTGGGTACACAGTGGATTTCA 943
QY 782 agtgcagagtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 837
Db 944 TGTGTAAGGTGATGAGTACCCGACCGACGACATCCAGTGGCTAAAGCAATCGAG 999

```

RESULT 9





Db	928	ATTCTGTGTGTCCTTCTTGAAAGGGCACTACACCTGCATCTGTGGAAATAGATGGGA	987
Qy	668	ccatcaacgcgcacctaacaagtgatgtgatccagcgcgacccgltccaaaccgcgtgtca	727
Db	988	GCATCAACCAACACCTACCACTGTGACGTCTGGAAAGATCTCCGACCCGACCATCTTC	1047
Qy	728	cagcgcacgaccccggttaaacagcagtgtagcttcggggggacacgltcccttcagttca	787
Db	1048	AGGcAGGgGcCTGScTGCcCAACAAACAGTGGCCCTGTGGGcAGcATGTGAGTTCATGTGTA	1107
Qy	788	aggtgcgcagcagcagcgttaagccggtgatccagtcggtctgaagcgcgttggag	837
Db	1108	AGGCTGTACAGCGATCCSMACCTTCACATTGATGCTGAAGACATCGAG	1157

RESULT 12  
 US-08-166-717D-1  
 ; Sequence 1, Application US/08166717D  
 ; Patent No. 5789182  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yayon, Avner  
 ; APPLICANT: Ornitz, David M.  
 ; APPLICANT: Klagsbrun, Michael  
 ; APPLICANT: Leder, Philip  
 ; TITLE OF INVENTION: SYSTEM FOR ASSAYING BINDING  
 ; TITLE OF INVENTION: TO A HEARIN-BINDING GROWTH  
 ; TITLE OF INVENTION: FACTOR RECEPTOR  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Clark & Elbing LLP  
 ; STREET: 176 Federal Street  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: U.S.A.  
 ; ZIP: 02110  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" diskette, 1.44 MB  
 ; COMPUTER: IBM COMPATIBLE  
 ; OPERATING SYSTEM: WINDOWS 95  
 ; SOFTWARE: wordperfect (Version 7.0)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/166,717D  
 ; FILING DATE: 12/14/93  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/631,717  
 ; FILING DATE: 12/20/90  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kristina Bleker-Brady  
 ; REGISTRATION NUMBER: 39,109  
 ; REFERENCE/DOCKET NUMBER: 00383/017002  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 723-4123  
 ; TELEFAX: (617) 723-8962  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3503  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-166-717D-1

	Query Match	3.18	Score 99.2	DB 1	Length 3503
	Best Local Similarity	50.28	Prod. No. 7	9e-14	
	Matches 326	Conservative	3	Mismatches 306	Indels 15
					Gaps
Qy	203	gagagcgcttcgcgcgtctcgcgcgaagggctgtgaagtgtgaacagatgtgagtcgaggaagt	262		
Db	508	GCACACCGTACCCGGCATYCCACAGGGGAGGAGGTGGAGGTGGCGGACATCCATCCCGCTGACT	567		
Yy	263	ccgagcgtgtacgtctgtgcgaagggccacaaagcgtcttcgcgcgaagcctgtagcgttcaactacacc	322		

[illegible]

RESULT 13  
 PCT-US93-05703-1  
 ; Sequence 1, Application PC/TUS9305703  
 ; GENERAL INFORMATION:  
 APPLICANT: Nova, Michael P.  
 APPLICANT: Gonzalez, Ana-Maria  
 APPLICANT: Baird, J. Andrew  
 TITLE OF INVENTION: Process for Detection of Neoplastic  
 TITLE OF INVENTION: Disease  
 NUMBER OF SEQUENCES: 2  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: Fitch, Even, Tabin & Flannery  
 STREET: 135 South LaSalle Street, Suite 900  
 CITY: Chicago  
 STATE: IL  
 COUNTRY: USA  
 ZIP: 60603  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US93/05703  
 FILING DATE: 19930614  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/900,646  
 FILING DATE: 18-JUN-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Seidman, Stephanie  
 REGISTRATION NUMBER: 33,779  
 REFERENCE/DOCKET NUMBER: 54625PCT



```

: TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
: NUMBER OF SEQUENCES: 18
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &
: ADDRESSEE: CUSHMAN
: STREET: 130 Water Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: US
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/471,570
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 435
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US/08/149,664
: FILING DATE:
: APPLICATION NUMBER: US 07/743369
: FILING DATE: 16-AUG-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: LINER, Ernest V
: REGISTRATION NUMBER: 29822
: REFERENCE/DOCKET NUMBER: 40897
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)523-3400
: TELEFAX: (617)523-6440
: TELEX: 200291 STRE UR
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1603 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 25..1602
: US-08-471-570-3

```

Query Match 2.8%; Score 88.2; DB 1; Length 1603;

Best Local Similarity 50.7%; Pred. No. 2e-11; Matches 244; Conservative 0; Mismatches 228; Indels 9; Gaps 1;

```

Oy 375 ctctcttggtggtcaagagagaccgcaagcagctggtgcaagcgcgttcacaca 434
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 93 CTCCTTCAGTTAGTTGAGGATACCACTTAGAGCCAGAGAGGAGCCACTGAGACCA 152
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 435 gccctccaaagtgaagcggtgtagcgaagcggtggtgtagcttcgagcgtcaa 494
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 153 CACAGAAAAGATGAAAAGCGGCTTCATGCTGTGCTCGGCAACACTGTCAATTTCG 212
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 495 gtgcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 554
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 213 CTGCCACGCGGGGGGAGAACCAATGCAACCATGCGGTGGCGTGAAGAAAGGAGATT 272
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 555 gaag-----ggtccagaagcggtgagcccaagaagaaggtgagactgagcct 605
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 273 TAAGCAGAGAGATCGCATGAGGCTTACAGGTACAGAAACAGCACTGAGCCTCATTTAT 332
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 606 gaagaacctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 665
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 333 GGAAGATGCTGCTCCATCTGACAAAGGAAATTTATCTGTGTGGTGGAGAAATGATACGG 392
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 666 cgcctcaacgcaacctcaaggtggtggtggtggtggtggtggtggtggtggtggt 725
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 393 GTCCATCAATCAACAGGTACCACTGATGTGTGTGAGGAGATGCGCTCAACGCGCATCTCT 452

```

```

Oy 726 caagagcagcagcccggtgaagcagcaggtgacttcgggggggaccacgctccagtg 785
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 453 CCAGCCGAGCTGCCGGAAATGCTCCACAGTGTGGAGAGACGTAGATTGTCTG 512
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 786 caaggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 845
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 513 CAAGTTTACAGTGTGCGCCAGCCGCGGACATCCAGTGTGATCAAGCAGTGAAGAAAGACGG 572
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 846 c 846
   |
Db 573 c 573

```

RESULT 16

```

: US-08-471-570-9
: Sequence 9, Application US/08471570
: Patent No. 5750371
: GENERAL INFORMATION:
: APPLICANT: IGARASHI, Koichi
: APPLICANT: SENOO, Masaharu
: APPLICANT: WATANABE, Tatsuya
: TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
: NUMBER OF SEQUENCES: 18
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &
: ADDRESS: CUSHMAN
: STREET: 130 Water Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: US
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/471,570
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/149,664
: FILING DATE:
: APPLICATION NUMBER: US 07/743369
: FILING DATE: 16-AUG-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: LINER, Ernest V
: REGISTRATION NUMBER: 29822
: REFERENCE/DOCKET NUMBER: 40897
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)523-3400
: TELEFAX: (617)523-6440
: TELEX: 200291 STRE UR
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2310 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 25..1980
: US-08-471-570-9

```

Query Match 2.8%; Score 88.2; DB 1; Length 2310;

Best Local Similarity 50.7%; Pred. No. 2.3e-11; Matches 244; Conservative 0; Mismatches 228; Indels 9; Gaps 1;

```

Oy 375 ctctcttggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 434
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 93 CTCCTTCAGTTAGTTGAGGATACCACTTAGAGCCAGAGAGGAGCACTACTGAGACCA 152

```

```

QY 435 gccctccaagatgagcgccggtgatcgacgcccgttggtacgtccgtgcgtctca 494
      |||||
Db 153 CACAGAAAGATGAAAGCGGCTCCATGCTGCTGGGCAACACGTCAAGTTG 212
QY 495 gtcgttgccagcgccaccctccgcccacatacgtgatgaagagccaccgctt 554
      |||||
Db 213 CTGCCCGCGGGGGGACCCCAATGCAACATGCGGTGGCTGAAAACGGAAGAGATT 272
QY 555 gacg-----cgccagagcgcttgaccagaaagtgacactagacct 605
      |||||
Db 273 TAAGCAGGAGCATTCGATTGAGGCTACAAAGTACGAACACGACTGAGCCTATTAT 332
QY 606 gaagaactcgcgcgcgaggaagacggcgcaatacacctcgcgcgttgtaaccgcgcg 665
      |||||
Db 333 GGAAGTGTGTCCTCCATCTGACAAAGGAAATATATCTGTGTGGAAGAAATATCCG 392
QY 666 cgcacatcaacgcacactacaagtgatgtgtatccagcgacccgttccaaagcgctgt 725
      |||||
Db 393 GTCATCAATTCACACTACCTACCTGATGTGTGGAGCGCATCGCCTCACCGGCCATCT 452
QY 726 caacagcagcagcccgctgaacacagcagcttgactcggggggaacacgctccactg 785
      |||||
Db 453 CCAAGCCGAGCTCCGCGAAATGCTCCACAGTGTGCGGAGGAGAGTAGAGTTTGTCTG 512
QY 786 caagtgccgagcagcgtgaagcgcgttgatccagtggtcgaagcggtgagtgacgcgc 845
      |||||
Db 513 CAAGGTTTACATGATGCCACGCCCAATCCATCCATGATCAACACACGTGGAAGAACGG 572
QY 846 c 846
Db 573 C 573

```

```

RESULT 17
US-08-070-165F-5
; Sequence 5, Application US/08070165F
; Patent No. 5750365
; GENERAL INFORMATION:
; APPLICANT: Chiu, Ing-Ming
; APPLICANT: Poulin, Matthew L
; TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ing-Ming Chiu
; STREET: 52052 Davis Medical Research Center, 480 West
; STREET: 9th Avenue
; CITY: Columbus
; STATE: Ohio
; COUNTRY: USA
; ZIP: 43210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/070,165F
; FILING DATE:
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (614)-293-8093
; TELEFAX: (614)-293-5631
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2675 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:

```

```

; ORGANISM: No. 5750365ophthalmus viridescens
; DEVELOPMENTAL STAGE: Adult
; TISSUE TYPE: Regenerating forelimb blastema
; CELL TYPE: Mesenchyme and Epithelium
; IMMEDIATE SOURCE:
; LIBRARY: lambda gt11
; CLONE: KP23-1
; POSITION IN GENOME:
; UNITS: bp
; FEATURE:
; NAME/KEY: 5' UTR
; LOCATION: 1..324
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 325..2511
; FEATURE:
; NAME/KEY: 3' UTR
; LOCATION: 2512..2675
; PUBLICATION INFORMATION:
; AUTHORS: Poulin, Matthew L
; TITLE: Nucleotide sequences of two new
; TITLE: factor receptor-2 variants
; JOURNAL: Biochim. Biophys. Acta
; VOLUME: 1220
; PAGES: 209-211
; DATE: 1994
; REFNAME: RESIDUES IN SEQ ID NO: 5: FROM 1 TO 2675
; US-08-070-165F-5

```

```

Query Match 2.7%; Score 87.2; DB 1; Length 2675;
Best Local Similarity 53.9%; Pred. No. 4,1e-11;
Matches 207; Conservative 0; Mismatches 168; Indels 9; Gaps 1;

```

```

QY 463 gcaagcccggtggttagctcgtcgcgtcaagtcgctgagcagcgagccctcgcc 522
      |||||
Db 556 GCTGTGCCCGCTGCCACACCTGTGAAGTCCGCTGCCAGCGGAGAACCTTACGCC 615
QY 523 gacatcagtgatgaagagcagcagcgcctgacg-----cgccagagcgct 573
      |||||
Db 616 TCCATGAGGTGCTGAGACGCAAGGAGTTCAAGCAGAGACCGCATTTGGCGCTTC 675
QY 574 gagccaggaagaagaagtgaacactgaagctgaagaactcgcgcgaggaagacagcg 633
      |||||
Db 676 AAGGTACGTAGTCAACACTTACGCTGATCATGAGAGCGGTGCTTCCCTGACGAGGC 735
QY 634 aaataacctgcgcgtgtcgaacgcgcgcgcgcacatcaacgcacactacaagtgat 693
      |||||
Db 736 AACTACACCTGTATCATGTGGAACGAGTATGATCCATCAATCACACCTACCTGAT 795
QY 694 gtgattcagcgagcccgcttccaaagcgcgtgctcaagcagcagcagcccgctgaacagc 753
      |||||
Db 796 GTTGTGAGACGGTCAACCCACCGGCCAATATCAAGCTGGGCTTCCGCAACACAC 855
QY 754 gtgacttcggggggaacagctccctcagtgcaagtgctgcaagcagcagcgtgagcggtg 813
      |||||
Db 856 ACAAAAGTTGGGGCGATGACAGATTGTTGCAAAAGTCTACAGTACGACGACACCAT 915
QY 814 atccagtgctgaagcgctgag 837
      |||||
Db 916 ATCCAATGATTCGACATTTTGAG 939

```

```

RESULT 18
US-08-885-418-5
; Sequence 5, Application US/08885418
; Patent No. 5925528
; GENERAL INFORMATION:
; APPLICANT: Chiu, Ing-Ming
; APPLICANT: Poulin, Matthew L
; TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)
; NUMBER OF SEQUENCES: 12

```



CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ing-Ming Chiu  
STREET: 62052 Davis Medical Research Center, 480 West  
STREET: 9th Avenue  
CITY: Columbus  
STATE: Ohio  
COUNTRY: USA  
ZIP: 43210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/885,418  
FILING DATE:  
CLASSIFICATION: 435  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (614)-293-8093  
TELEFAX: (614)-293-5631  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2675 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: NO. 5925528ophthalmus viridescens  
DEVELOPMENTAL STAGE: Adult  
TISSUE TYPE: Regenerating forelimb blastema  
CELL TYPE: Mesenchyme and Epithelium  
IMMEDIATE SOURCE:  
LIBRARY: lambda gt11  
CLONE: kp23-1  
POSITION IN GENOME:  
UNITS: bp  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..324  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 325..2511  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 2512..2675  
PUBLICATION INFORMATION:  
AUTHORS: Poulin, Matthew L  
TITLE: Nucleotide sequences of two new  
TITLE: (No. 5925528ophthalmus viridescens) fibroblast growth  
JOURNAL: Blochim. Biophys. Acta  
VOLUME: 1220  
PAGES: 209-211  
DATE: 1994  
RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 TO 2675  
US-08-885-418-5

Query Match	2.7%	Score 87.2	DB 2	Length 2675;
Best Local Similarity	53.9%	Pred. No. 4,1e-11;		
Matches	207;	Conservative	0;	Mismatches 166; Indels 9; Gaps 1;
QY	463	gcaacgcccgtggtagctccgttcgctcgaatgctgtggtgccagcgcgacccctcgccc	522	
Db	556	gctgtgccccgcgtcccaacactgtgaagttccgccttccacgcggtggcgaaccccttgcggcc	615	
QY	523	gacatcagcttgatgaagacgcagcaccgaccttgcg-----cgccacgaagcgct	573	
Db	616	tcctatgagctggtctcaaacacggcgaagagttcgaagcagacaccgacatttgccgcgttc	675	

```

Oy 574 gagcccaagaagaagatggaacactagcctgaagaacctgcgcggaggaacagcgc 6333
Db 676 AAGGTACCTTAATCAACACTTTCAGCTTATGATGAGAGAGCTGGTCCCTGTGACAGAGGC 7355
Oy 634 aatacacactgcgcgcgltgcgaaccgcgcgcgcgcacatacgaaccctacaagttgat 6933
Db 736 AAGTACACAGCTTATCATGAGAGAACGAGATGATTCATATCATACACTTACCACTGAGAT 7955
Oy 694 gtatcacagcggaaccggttcccaagcccgltgtcacagagcaagcaaccccgtagaacaagc 7533
Db 796 GTTGTGCGAGCGGTACACCCACCGGCAATATCTCCAAACGTGGGCTTCGGGAAACACAAAC 8555
Oy 754 gtggacttcgcggggggaacacgctcccttcagtgcaaggtgcgcgcgcagcgtgaagcgcgtg 8133
Db 856 ACAAAAGTTGGGGGGCGATGCAAGAGTTGTTTTCGCAAAAGTCTACAGTGAACGACACAGCCAT 9155
Oy 814 atccagtgctcgaagcgcggtgag 837
Db 916 ATCCATGATTCGACATTTTGAG 939

RESULT 19
US-08-070-165F-9
; Sequence 9, Application US/08070165F
; Patent No. 5750365
GENERAL INFORMATION:
APPLICANT: Chiu, Ing-Ming
APPLICANT: Poulin, Matthew L
TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ing-Ming Chiu
STREET: 82052 Davis Medical Research Center, 480 West
STREET: 9th Avenue
CITY: Columbus
STATE: Ohio
COUNTRY: USA
ZIP: 43210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/070.165F
FILING DATE:
CLASSIFICATION: 435
TELECOMMUNICATION INFORMATION:
TELEPHONE: (614)-293-8093
TELEFAX: (614)-293-5631
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2681 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: No. 5750365ophthalmus viridescens
DEVELOPMENTAL STAGE: Adult
TISSUE TYPE: Regenerating forelimb blastema
CELL TYPE: Mesenchyme and Epithelium
IMMEDIATE SOURCE:
LIBRARY: lambda gt11
CLONE: KP19-1
POSITION IN GENOME:
UNITS: bp
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..324

```

RESULT 19  
 US-08-070-165F-9  
 Sequence 9, Application US/08070165F  
 Patent No. 5750365  
 GENERAL INFORMATION:  
 APPLICANT: Chiu, Ing-Ming  
 APPLICANT: Poulin, Matthew L  
 TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)  
 NUMBER OF SEQUENCES: 12  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: Ing-Ming Chiu  
 STREET: 52052 Davis Medical Research Center, 480 West  
 STREET: 9th Avenue  
 CITY: Columbus  
 STATE: Ohio  
 COUNTRY: USA  
 ZIP: 43210  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/070,165F  
 FILING DATE:  
 CLASSIFICATION: 435  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (614)-293-8093  
 TELEFAX: (614)-293-5631  
 INFORMATION FOR SEQ ID NO: 9:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2681 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: NO. 5750365ophthalmus viridescens  
 DEVELOPMENTAL STAGE: Adult  
 TISSUE TYPE: Regenerating forelimb blastema  
 CELL TYPE: Mesenchyme and Epithelium  
 IMMEDIATE SOURCE:  
 LIBRARY: lambda gtl11  
 CLONE: KP19-1  
 POSITION IN GENOME:  
 UNITS: bp  
 FEATURE:  
 NAME/KEY: 5'UTR  
 LOCATION: 1..324  
 FEATURE:









RESULT 26  
US-08-690-473-1  
; Sequence 1, Application US/08690473  
; Patent No. 5876923  
; GENERAL INFORMATION:  
; APPLICANT: Leopardi, Rosario  
; APPLICANT: Roizman, Bernard  
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 AS AN  
; TITLE OF INVENTION: INHIBITOR OF APOPTOSIS  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P. O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/690.473  
FILING DATE: 26-JUL-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: ARCD:239  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4257 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-690-473-1

Query Match 1.7%; Score 54.4; DB 2; Length 4257;  
Best Local Similarity 44.3%; Pred. No. 0.0017;  
Matches 265; Conservative 0; Mismatches 331; Indels 2; Gaps 1;  
Db 757 gacttcgggggagacacgctctccacgtgtaaggtgacgagcgtgtaagcgggtgac 816  
1735 GACCGGCGCGCAGAGGAGCTTCTGACACGCTGCGCGCGCTTACGCGCCCTGTTG 1794  
Qy 817 cagtgtgtgaagcgctgtgagtgacgagcgagcgagcgacacacacacacgtgtg 876  
1795 GCGCGCAGAACCGCGCGCTGACGGGGGCGCGGGGAGCCCGCGCGCGAGATGAC 1854  
Qy 877 ggcgcgcagaagttgt 936  
1855 GAGGGGGTCCCGGTAC 1914  
Qy 937 tacctcaataagctgtctatcatcacccgtgtgtgtgtgtgtgtgtgtgtgtgtgt 996  
1915 GGCCTCCCGGGGATCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1974  
Qy 997 ctgt 1056  
1975 GGGGGCAGACCG 2034  
Qy 1057 ccaaaacgcgcagagcgacacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1116  
2035 GCCAGGCG 2094  
Qy 1117 gt 1176  
2095 CTGGCCGAGGAGCTTCTGACGGGACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2154

RESULT 27  
US-09-259-821A-1  
; Sequence 1, Application US/09259821A  
; Patent No. 6210926  
; GENERAL INFORMATION:  
; APPLICANT: LEOPARDI, ROSARIO  
; APPLICANT: ROIZMAN, BERNARD  
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS  
; FILE REFERENCE: ARCD:317  
; CURRENT FILING DATE: 1999-03-01  
; PRIOR APPLICATION NUMBER: 08/690.473  
; PRIOR FILING DATE: 1996-07-26  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4257  
; TYPE: DNA  
; ORGANISM: HERPES VIRUS, TYPE 1  
US-09-259-821A-1

Query Match 1.7%; Score 54.4; DB 4; Length 4257;  
Best Local Similarity 44.3%; Pred. No. 0.0017;  
Matches 265; Conservative 0; Mismatches 331; Indels 2; Gaps 1;  
Qy 757 gacttcgggggagacacgctctccacgtgtaaggtgacgagcgtgtaagcgggtgac 816  
1735 gacgcgcgcagaaaggtctctctgtgtacacgctgtgtgtgtgtgtgtgtgtgtgt 1794  
Qy 817 cagtgtgtgaagcgctgtgagtgacgagcgagcgagcgacacacacacacgtgtg 876  
1795 ggcgcgcagaagcgctgt 1854  
Qy 877 ggcgcgcagaagttgt 936  
1855 ggggggt 1914  
Qy 937 tacctcaataagctgtctatcacccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 996  
1915 ggcgcgcgcagaaaggt 1974  
Qy 997 ctgt 1056  
1975 ggggggcgcagacccgc 2034  
Qy 1057 ccaaaacgcgcagagcgacacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1116  
2035 ggc 2094  
Qy 1117 gt 1176  
2095 ctgt 2154  
Qy 1177 tgc 1236  
2155 gccagcccccgc 2212  
Qy 1237 ccgc 1296

























---

THIS PAGE IS BLANK



```

153 etgluAspProProtyrSerThrSerTyrrArgAlaProPheTrpSer 169
170 GlnProGlnMetLysLysLysLeuTyrrAlaValProAlaGlyAsnThr 186
483 CGTGGCGCTAAGTGTGGCGAGCGGCGCGCTCGGCGCGCATCATGCT 532
186 rValysPheArgCysProSerAlaGlyAsnProThrProGlyTyrLeuArg 203
533 GGATGAAGAGCAGCAGCGCTTG.....ACGCGCCGAGAGCGCGCT 573
203 rPheuLysAsnGlyArgGluPheGlyGlyLysIleHisArgIleGlyIle 219
574 GAGCCCGAGAGAGAGAGTGGACACTGAGCTGAGAGAACTGGCGCGGA 623
220 ArgLeuArgGlnIstGlnIstTrpSerLeuValMetGluSerValValProSe 236
624 GAGCAGCGGGAATATACACTGCGCGCTGTGAGAACCGCGCGCGCATCA 673
236 rAspArgGlyAsnTyrrThrCysLeuValGluAsnLysPheGlySerIleS 253
674 AGCGCCACTACAGAGTGTGATGTATCCAGCGGAGCCGCTCCAAAGCCGTG 723
253 eTyrrSerTyrrLeuAspValLeuGluArgSerProHisArgProIle 269
724 CTGACAGCGAGCGCGCGCGTGAACAGCAGCGAGTGTGAGTGGGGGAGAC 773
270 LeuIleValGlyLeuProAlaAsnThrThrAlaMetLeuGlySerAspVa 286
774 GTCTCTTCAGTGTGAAGTGTGGCGAGCGAGTGAAGCCGCGTATCCAGTGGC 823
286 IglrPhePheCysLysValTyrrSerAspAlaGlnProHisIleGlnTrpL 303
824 TGAAGCGCGTGGAG.....TACGCGCGCGAGCGCGCGCAC 858
303 euLysHisIleGluValAsnGlySerArgTyrrGlyProAsp..... 316
859 AACGCCACATCGATGTGGCGCGCAGAGATTGTG...GTGGTGGCCAC 905
317 .....GlyValProPheValGlnValLeuLysThr 326
906 GGGTGTGAGTGTGTGGCGCGCGAGCGCTCTTACCTCAATAGCTGTCTCA 955
326 rAlaAspIle.....AsnSerSerGluValGluValLeuTyrrL 339
956 TCACCCGTGCGCGCGAGAGATCGGGCATGTACATCTGCTTGGCGCC 1005
339 euHisAsnValSerPheGluAspAlaGlyLysTyrrThrCysLeuAlaGly 355
1006 AACCCAGTGGGCTACAGTTCGCGAGCGCGCTTCTCACCGGCTGGCGCAGA 1055
356 AsnSerIleGlyLeuSerTyrrGlnSerAlaTrpLeuThrValLeuProGln 372
1056 C.....CCAAACCGCGAGCGCGACCTGTGCGCTCTCTGTCT 1093
372 uGluAspPheAlaLysGluAlaGlnGlyProGluThrArgTyrrThrAsp 388
1094 CGGCGCTAGCTGCGCTGCGCGCGCTGCTCATCGGCATCCAGCGCGCGCT 1143
389 .....IleIleIleTyrrThrSerGlySerLeu 397
1144 GTCTTCATCTGGGCGAGCTGTCTGTGCTGTGGCTGGCAGCGCGCAGAGAA 1193
398 AlaLeuLeuMetAlaAlaValIleValValLeuCysArgMetGln..... 412
1194 GCGGTGACCGCGCGCGCTGCGCTCTCTGCGCTGTGGCGAGCGCGCGCG 1243
413 .....LeuProProThrLysThrHisLeuGluProAla 423
1244 GAGCGCGCGCGAGCGCGAGAGACAGCAAGACTTCCTCTGCTTGGCGCC 1293

```

```

423 lathrValHisLysLeuSer.....ArgPheProLeuMetArgGln 436
1294 CTCAGCGCTGGCGCTGTGTGGCGCTGTGTGAGAGCATGAGTCTCGCGG 1343
437 PheSerLeu.....GluSerSerSerSerI 445
1344 AGCCCCCGACACTTACTGCGCGCGAGCGCGCGAGTGTGCGCGCTTGT 1393
445 y.LysSerSerThrSerLeuValArgValThrArgLeuSerSerSerCys 461
1394 ACCCCA 1399
462 ThrPro 463
seq_name: p1r2:S19947

```

```

seq_documentation_block:
  fibroblast growth factor receptor - Iberian ribbed newt
  C:Species: Pleurodeles waltl (Iberian ribbed newt)
  C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Jul-1999
  C:Accession: S19947
  R:Shi, D.L.; Feige, J.J.; Riou, J.F.; Desimone, D.W.; Boucay, J.C.
  submitted to the EMBL Data Library, March 1992
  A:Description: Receptors during early development of the urodele Pleurodeles waltl
  A:Reference number: S19947
  A:Accession: S19947
  A:Status: preliminary
  A:Molecule type: mRNA
  A:Residues: 1-822 <SH1>
  A:Cross-references: EMBL:X65059
  C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; pr
  C:Keywords: ATP; growth factor receptor
  F:283-354/Domain: immunoglobulin homology <IMM>
  F:484-769/Domain: protein kinase homology <KIN>
  F:492-500/Region: protein kinase ATP-binding motif

```

```

alignment_scores:
  Quality: 508.00      Length: 486
  Ratio: 1.801        Gaps: 19
  Percent Similarity: 58.025  Percent Identity: 29.012

```

```

alignment_block:
  US-09-598-042a-2 x S19947 ..

```

```

Align seg 1/1 to: S19947 from: 1 to: 822

```

```

19 CGATGTCTCCCGCGCGCATGCGCGCGGAGTTGCTTCATC..... 60
36 ArgThrLeuProGlnGlyArgLysAlaAsnLeuAlaGlnLeuValSerGln 52
61 .....CACGTATGTCTGTCGCGGCGTGCAGCATTTCTGCGCAGG 100
52 uGluGlnGlnHisPheLeuLeuAspPro..... 61
101 TGCGCGCGGTGGCGCGCATATGCGGCTGCATGCGCCAGTGAAGGGAGAC 150
62 .....GlyAsnAlaLeuArgLeuPheCys.....AspTrpAsn 72
151 CCGCGCGCGCTGACCATGTGG...ACCAAGATGGCGCGCATCCACAG 197
73 GlnThrThrIleValAsnTrpTyrrThrGlnSerThrArgLeuGlnHisGln 89
198 CGGCTGAGCGCGCTTCGCGCTGTGCGCGCAGGGGCTGAAGTGAAGCAGG 247
89 yGly.....ArgIleArgLeuThrAspThrValGlnGlnIleAlaAspV 104
248 TGAAGCGGAGGATGCGCGGCTGTACGTGTGAAGCGCCACCAAGCGCTTC 297
104 alThrTyrrGluAspSerGlyLeuTyrrLeuCys...ValValProGlyThr 119
298 GGCAGCTGTGAGCGTCAACTACACCTTCGTGCTGTGATACATTAAGCC 347

```

```

120 GlnHisIleLeuArgAspPheThrIleSerValValAspSerLeuAlaSe 136
348 AGCG.....AAGAGACCTGGGGCCGACAGCTCTCTGGG... 384
136 rGlyAspAspAspAspGlyAspHisGlyArgGlyAspSerAlaGlyAspM 153
385 ..GGTCAAGAGACCCCGCCAGCCAGCAGCGGGGCGACCCCGCTTCA 432
153 etGlyGlyAspProProTyrSerThrSerTyrAlaGlyAlaProPheTrpSer 169
433 CAGCCCTTCAAGATGAGCGCGCGGTATGCGACGCGCCGTGGGTAGCTC 482
170 GlnProGlnArgMetAspIlySerIleuTyrAlaValProAlaGlyAsnTh 186
483 CGTGGCGGTCAAGTGGCGTGGCGGACGCGACCCCTGGCCCAATCAGCT 532
186 rValIlySerPheArgCysProSerAlaGlyAsnProThrProGlyIleArgT 203
533 GGATGAAGACGACGACGCTTG.....ACGCGCCGACGAGCGCGCT 573
203 rPheUlyAsnGlyArgGlyPheGlyGlyGlnHisArgTlleGlyGlyTlle 219
574 GAGCCGAGGAAGAAGAACTGACATGACCTGAGAACTGCGGCGCGA 623
220 ArgLeuArgHisGlnHisIleThrSerIleuValMetGlySerValAlaProSe 236
624 GACAGCGCGCAATACATACCTCCCGGTGTGACACCGCGCGCGCCATCA 673
236 rAspArgGlyAsnTyrThrCysIleuValGlyAsnIlySerPheGlySerIleS 253
674 ACGGCACCTACAGGTGATGTATCCAGCGGACCCGCTTCCAGACCCCTG 723
253 eTyrTyrSerTyrLeuLeuAspValLeuGlnArgSerProHisArgProIle 269
724 CTCACAGGACGACGACCCCGTGAACAGACGAGTGCATCTGGGGGACGAC 773
270 LeuGlnAlaGlyLeuProAlaAsnThrThrAlaMetLeuGlySerAspVal 286
774 GTCTCTTCAGTGCAGAGTGGCGGACGACGATGAAAGCGGTATCCAGTGGC 823
286 IGlPhePheCysIlySerValTyrSerAspAlaGlnProHisIleGlnTrpL 303
824 TGAAGCGCGGAGAG.....TACGGCGCGGAGCGCGCGCAC 858
303 euLysHisIleGlyValAsnGlySerArgTyrGlyProAsp..... 316
859 AACTCCACCATGATGTGGCGGCGCAGAAATTGTG...GTGCTGCCAC 905
317 .....GlyValProPheValGlnValLeuLysTh 326
906 GGTGACGTGTGTGTCGCGCGCGACGCTCTTACCTCAATAAGCTGTCA 955
326 rAlaAspIle.....AsnSerSerGlyValGlyValLeuTyrL 339
956 TCACCCGCGCGCGCGACGATGCGGCGATGTACATGTGCTTGGCGGCC 1005
339 euHisAsnValSerPheGlyAspAlaGlyGlyTyrThrCysLeuAlaGly 355
1006 AACACCATGAGCTTACAGCTTCCGACGCTTCTTCTCACGCTGTGCGCA 1055
356 AsnSerIleGlyLeuSerTyrGlnSerAlaTrpLeuThrValLeuProGln 372
1056 C.....CCMAAACGCGACGCGCACCTGTGGCTCTCTGTCTCT 1093
372 uGlyAspPheAlaLysGlnAlaGlyGlyProGlnThrArgTyrThrAsp. 388
1094 CGGCACATAGCTGCGCGCGCGGTGATCGGATCCACGCGCGCT 1143
389 .....IleIleIleTyrThrSerGlySerLeu 397
1144 GTCTTCATCTGGGACACCTGCTCTGTGCTTGGCTTGGACGCGCGCA 1193
398 AlaLeuMetAlaAlaValIleValValLeuCysArgMetGln.... 412

```

```

1194 GCGTGCACACCCCGCGCTGCCCTCCCTGCTGGGACGCGCGCGG 1243
413 .....LeuProThrThrIleThrIleGlnIleProA 423
1244 GACGCGCGCGGACCGACGCGGACAGACCTTCCCTGTTGGCGCGC 1293
423 lathValHisIlySerLeu.....ArgPheProLeuMetAlaGln 436
1294 CTCAGCGCTGGCGCTGTGTGGGCGCTGTGTAGAGCATGGCTCCGCGC 1343
437 PheSerLeu.....GlySerSerSerGln 445
1344 AGCCCCCAGACTTACTGGGCGCGCGCGCGCTTCCGCTTACTGTCT 1393
445 Y.LysSerSerThrSerLeuValArgValThrArgLeuSerSerCys 461
1394 ACCCA 1399
462 ThrPro 463
seq_name: p1r2:A48991
seq_documentation_block:
heparin-binding growth factor receptor - mouse
N;Alternate names: HBGF receptor
N;Contains: protein-tyrosine kinase (EC 2.7.1.112)
C;Species: Mus musculus (house mouse)
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 04-Feb-2000
C;Accession: A48991
R;Kato, O., Hatori, Y., Sasaki, H., Sakamoto, H., Fujimoto, K., Fujii, T., Sugimura
Cancer Res. 53, 1136-1141, 1993
A;Title: Isolation of the complementary DNA encoding a mouse heparin-binding growth f
A;Reference number: A48991, MIMD:93177694
A;Accession: A48991
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1800 <KAT>
A;Cross-references: GB:556291; NID:g298329; PIDN:AAB25535.1; PID:g298330
A;Experimental source: brain
A;Note: sequence extracted from NCBI backbone (NCBIN:126536, NCBI:P:126537)
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; pr
C;Keywords: hrgp; growth factor receptor; heparin binding; phosphotransferase; tyrosin
F;262-335/Domain: immunoglobulin homology <IMH>
F;464-749/Domain: protein kinase homology <KIN>
F;472-480/Region: protein kinase ATP-binding motif
alignment_scores:
Quality: 490.00 Length: 470
Ratio: 1.750 Gaps: 18
Percent Similarity: 59.574 Percent Identity: 32.128
alignment_block:
US-09-598-042A-2 x A48991 ..
Align seg 1/1 to: A48991 from: 1 to: 800
9 GAGCCGACGTGCATGCTCCGCGCGCATGGAGCGCGGATGTGCTTCA 58
||| |||||..... |||||... ||| |||||:
17 GAlaIatThrSerGlnProGlyProGlyGlnArg.....ValValAr 31
||| |||||..... |||||..... |||||.....
59 TCCACGTATGCTGTCCCGGCGTGCAGATTTCTGCGCAG..... 99
||| |||||..... |||||..... |||||.....
31 gArgAlaIatGlnValProGlyProGlyProSer.GlnGlnGlnVal 47
||| |||||..... |||||..... |||||.....
100 GTGGCGCGCGTGGCGCGCGCATATGCGGCTGCAGTCC...CCAGTGAAGG 146
||| |||||..... |||||..... |||||.....
48 AlaPheGlySerGlyAspThrValGlnIleuSerCysHisProProGly 64
||| |||||..... |||||..... |||||.....
147 GAGCCGCGCGCGCGCATGATGTGACCAAGATGAGCGCGCATGCCA 196
||| |||||..... |||||..... |||||.....
64 yAlaProThrGlyProThrValTrpAlaIlyAspGlyThrGlyLeuVal 81

```

```

197 GCGGCTGAGACCCGCTTCGCGCTGCTGCGGAGGCGCTGAAGCTGAAGCAG 246
    :      :      :      :      :      :      :      :      :
81  laser...HisArgIleLeuValGlyProGlnArgLeuGlnValLeuAsn 96
247 GTGAGGCGGAGAGTGGCGCGGTGACGTGTC.....AAGGCCACCA 290
    :      :      :      :      :      :      :      :      :
97  AlSerHisGlnAspAlaGlyValTyrSerCysGlnHisArgLeuThr 113
    :      :      :      :      :      :      :      :      :
291 CGGCTTCGCGACGCTGACGCTCAACCTACACCTGTCGTCGTGAATGACA 340
    :      :      :      :      :      :      :      :      :
113  g.....ArgValLeuCysHisPheSerValArgValThrAspLar 127
341 TTAGCCCGAGGAGAGAGAGCGCTGGGCGCCGACAGCTTCCTGGGGGCA 390
    :      :      :      :      :      :      :      :      :
127  roSerSerGlyAspAspGlyAspGlyAspValAla..... 139
391 GAGAGACCCCGACAGCCAGAGTGGGACGAGCCGCTTCACAGACCCCTC 440
    :      :      :      :      :      :      :      :      :
140  GlnAspThrGlyAlaProTyrTrp.....ThrArgProGln 151
441 CAAGATGAGCGCGCGGTGATCGACGCGCGCTGGTACCTCGTGCGGC 490
    :      :      :      :      :      :      :      :      :
151  uArgMetAspLysLysLeuLeuAlaValProAlaAlaAsnThrValArgp 168
491 TCAGTGTGCGTGGCGGCGGACCGCTGCGCCGACATCAGTGGATGAG 540
    :      :      :      :      :      :      :      :      :
168  heArgCysProAlaAlaGlyAsnProThrProSerIleSerThrLeuLys 184
541 GAGACCGAGCGCTTGACG.....CGCCGAGGCGCGCTGAGCCGAC 581
    :      :      :      :      :      :      :      :      :
185  AsnGlyLysGlnPheArgGlyGlnHisArgIleGlyIleLysLeuArg 201
582 GAAGAAGAAGTGGACATGAGCCGGAAGAAGCTCGCGCGGAGACAGCG 631
    :      :      :      :      :      :      :      :      :
201  gHisGlnGlnTrpSerLeuValMetGluSerValValProSerAspArg 218
632 GCAATGACACCTCGCGCGGTGTCGAACGCGCGCGCGCATCAACGACAC 681
    :      :      :      :      :      :      :      :      :
218  LysnTyrThrCysValValGlnLysnLysPheGlySerIleArgGlnThr 234
682 TACAAGGTGATGTGATTCACGCGGACCGCTTCACAGCCGCTGTCACAG 731
    :      :      :      :      :      :      :      :      :
235  TyrThrLeuAspValLeuGlnArgSerProHisArgProIleLeuGln 251
732 CAGCCACCGCTGAAACAGACGAGTGGAGCTTGGGGGACACAGCTCTCC 781
    :      :      :      :      :      :      :      :      :
251  agLysLeuProAlaAsnGlnThrAlaIleLeuGlySerAspValGlnPhe 268
782 AGTCAAGGTGCGGACGAGCAGCTGAAGCGGTGATTCAGTGGCTGAAGCG 831
    :      :      :      :      :      :      :      :      :
268  IsCysLysValTyrSerAspAlaGlnProHisIleGlnTrpLeuLysHis 284
832 GTGAGTAGCGCGCGGAGCGCGCCACAACTCCACCATCGATGGGCGG 881
    :      :      :      :      :      :      :      :      :
285  ValGlnVal.....AsnGlySerLysValGlyProAspGln 296
882 CCAGAAGTTGTG...GTGCTGCCACGCGGTGAGCTGTGGCGGGCGG 928
    :      :      :      :      :      :      :      :      :
296  yThrProTyrValThrValLeuLysTrpAlaGly.....AlaAsnThr 311
929 ACGGCTCTCACTCAATAGCTCATACCGCTGCGCGCCGACAGAGAT 978
    :      :      :      :      :      :      :      :      :
311  hrAspLysGlnLeuGlnValLeuSerLeuHisAsnValThrPheGlnAsp 327
979 GCGGCGATGATCATCTGCTTGGCGGCAACACATGGGCTACAGCTTCG 1028
    :      :      :      :      :      :      :      :      :
328  AlaGlyLysTyrThrCysLeuAlaGlyAsnSerIleGlyPheSerHis 344
1029 CAGGCGCTCTCACTCACTGCTGCGCA...GACCCAAACCGCAGGCGCAC 1075
    :      :      :      :      :      :      :      :      :
344  sSerAlaTrpLeuValValLeuProAlaGlnGlnLeuMetGlnThr 361

```

```

1076 CTGTGACCTCTCTGTCCTCGGCCACTAGCTGCGTGGCGCGGTGATC 1125
    :      :      :      :      :      :      :      :      :
361  spGlnAlaGlySerValTyrAlaGlyValLeuSerTyrGlyValAlaPhe 377
1126 GGCATCCAGCGCGCGCTGTTTCATCTGGGACCGCTGCTCTGTGCT 1175
    :      :      :      :      :      :      :      :      :
378  .....PheLeuPheIleLeuValValAlaAlaValIleLe 389
1176 TTGCCGCGCCAGAGAAAGCGGTGCACCGCGCGCTGCGCCCTCCG 1224
    :      :      :      :      :      :      :      :      :
389  uCysArgLeuArgSerProProLysLysGlyLeuGlySerProThrValn 406
1225 .....CCTGGGACCGCGCGCGGAGCGCGCGCGAC 1257
    :      :      :      :      :      :      :      :      :
406  IsLysValSerArgPheProLeuLysArgGlnValSerLeuGlnSerAsn 422
1258 CGCAGCGGAGACAGAGACCTTCCC.....TGTGTGGCGCGCTCAGCC 1301
    :      :      :      :      :      :      :      :      :
423  SerSerMetLysSerAsnThrProLeuValArgIleAlaArgLysSer 439
1302 TTGCCCTGTGTTGGGCGTGTGTT.....GAGGACATGGGTCTCCGCGAG 1345
    :      :      :      :      :      :      :      :      :
439  rGlyGlnGlyProValLeuAlaAsnValSerGlnLeuGlnLeuProAla 456
1346 CCCCCCAG 1353
    :      :      :
456  sProLys 458

```

```

seq_name: pir2:155363
seq_documentation_block:
  fibroblast growth factor receptor 3 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 16-Jul-1999
C:Accession: 155363; B53627
R:Ornitz, D.M.; Leder, P.
J. Biol. Chem. 267, 16305-16311, 1992
A:Title: Ligand specificity and heparin dependence of fibroblast growth factor recept
A:Reference number: 155363; MUID:92355991
A:Accession: 155363
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-801 <RES>
A:Cross-references: GB:M01342; NID:g199144; PID:AA93935.1; PID:g199145
R:Chellatah, A.T.; McEwen, D.G.; Werner, S.; Xu, J.; Ornitz, D.M.
J. Biol. Chem. 269, 11620-11627, 1994
A:Title: Fibroblast growth factor receptor (FGFR) 3. Alternative splicing in immunogl
A:Reference number: A53627; MUID:94209351
A:Accession: B53627
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 242-364 <CHE>
A:Cross-references: GB:L26492
C:Genetics:
A:Gene: mfr3
A:Introns: 304/3; 353/1
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; pr
C:Keywords: ATP; growth factor receptor
F:262-335/Domain: immunoglobulin homology <IMM>
F:464-750/Domain: protein kinase homology <KIN>
F:472-480/Region: protein kinase ATP-binding motif

```

```

alignment_scores:
  Quality: 490.00      Length: 470
              Ratio: 1.750      Gaps: 18
Percent Similarity: 59.574      Percent Identity: 32.128

```

```

alignment_block:
  us-09-598-042a-2 x 155363

```

```

Align seq 1/1 to: 155363 from: 1 to: 801

```

```

9  GAGCCGACGTCATGCTCCGCGCCGATGAGCCGGGATTTGCTTCA 58
   ||| |||||::: ||||| :: ||| |||||:::
17  GAlaThrSerGluProGlyProGluGlnArg.....ValValAr 31
   : : : : :
59  TTCACGTCGATGCTGTTCCGGGTCGACGATTTTCGTCGACG 99
   : : : : :
31  gArgAlaIaGluValProGlyProGluProSer. GlnGlnGluVal 47
   : : : : :
100 GTGCGCCGCGTCGGCCGACATGCGCGTCGACGTC...CCAGTGGAGG 146
   : : : : :
48  AAlaPheGlySerGlyAspThrValGluLeuSerCysHisProGlyG 64
   : : : : :
147 GGACCCGCGCGCGTCGACGATGTCGACGACGATGGCGCGCACACCA 196
   | ||| |||||::: ||||| |||||
64  YAlaProThrGlyProThrValITrPAlaLysAspGlyThrGlyLeuVal 81
   : : : : :
197 GCGGTCGAGGCGCTTCCGCGTCGTCGCGCAGGCGGTCGAAGTGAAGCAG 246
   : : : : :
81  IAser...HisArgIleLeuValGlyProGlnArgLeuGlnValLeuAsn 96
   : : : : :
247 GTGAGCGCGGAGATGCCGCGCTGACGTCGTC...AAGCCACCA 290
   : : : : :
97  AAlaSerHisGluAspAlaGlyValIleSerCysGlnHisArgLeuThrAr 113
   : : : : :
291 CGGCTTGGGACGCTGACGCTCACTACACCCCTGCTGCTGATGACA 340
   : : : : :
113 9.....ArgValLeuCysHisPheSerValArgValIThrAspAla 127
   : : : : :
341 TTAGCCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 390
   ||| ||| : : : : : ||| ||| : : :
127 roserSerGlyAspAspGluAspGlyGluAspValAla..... 139
   : : : : :
391 GAGAGCCCCGACAGCCAGCAGTGGGCGACAGCCGCTTACACACCCCTC 440
   ||||| : : : : : ||| : : : : :
140 GIlusPrThrGlyAlaProIleTrp.....ThrArgProG 151
   : : : : :
441 CAAGATGAGGCGCGGTCGATGTCGACGCGCGTGGTACGTCGCGGCGG 490
   : : : : :
151 uArgMetAspLysLysLeuLeuAlaValProAlaIaAsnThrValAr 168
   : : : : :
491 TCAAGTCGTCGTCGACGAGCGGACCCCTGCGCCGACATCAGTCGATGAAG 540
   : : : : :
168 heArgCysProAlaIaGlyAsnProThrProSerIleSerIThrLeuLys 184
   : : : : :
541 GACGACCGCCCTGACG.....CGCCACAGCGCGCTGAGCCAG 581
   : : : : :
185 AsnGlyLysGluPheArgGlyGluHisArgIleGlyIleLysLeuAr 201
   : : : : :
582 GAGAGAGAGTGGACATGACGCTGAGACACCTGCGCGCGGAGGACAGC 631
   | : : : : : ||||| : : : : :
201 gHisGlnIThrSerLeuValMetGluSerValAlaProSerAspArg 218
   : : : : :
632 GCAATATACCTGCGCGCTGTCGAAACCGCGGCGGCGCATCAACCCAC 681
   ||||| : : : : : ||||| : : : : :
218 LysnITrThrCysValValGluAsnLysPheGlySerIleArgGlnThr 234
   : : : : :
682 TACAAGTCGATGTCATCCAGCGGACCCGTCGAAAGCCGTCGTCAGAG 731
   ||| : : : : : ||||| : : : : :
235 TTrIThrLeuAspValLeuGlnArgSerIThrnIThrProIleLeuGln 251
   : : : : :
732 CAGGACCCGCGTGAACAGACGATGTCGTCGTCGTCGTCGTCGTCGTC 781
   ||||| : : : : : ||||| : : : : :
251 hGlyLeuProAlaAsnGlnIThrAlaIleLeuGlySerAspAlaGluPhe 268
   : : : : :
782 AGTGAAGTCGTCGACGACGATGAGAGCCGTCGATGTCGTCGTCGTCG 831
   : : : : :
268 IAsCysLysValITrSerAspAlaGlnProHisIleGlnITrPheLysHis 284
   : : : : :
832 GTGAGTCGTCGCGCGGAGCGCGGACACACTGTCACCATGATGTCGCGG 881
   ||||| : : : : : ||||| : : : : :
285 ValGluVal.....AsnGlySerLysValGlyProAspG 296
   : : : : :
882 CCAGAGTTTGTG...GTGTCGCGCCAGGCTGACCTGTGTGCGCGCGCG 928

```

```

1  : : : : : ||||| ||||| : : : : :
296 YThrProGlyValIThrValLeuLysIThrAlaGly.....AlaAsnIThr 311
   : : : : :
929 ACGGCTCTACCTCAATAAGATGTGTATACACCGTCGCGCGGACGACGAT 978
   : : : : :
311 hAspLysGluLeuGlnValLeuSerLeuHisAsnValIThrPheGlnAsp 327
   : : : : :
979 GCGGCGATGTCATGCTGCTTGGCGCGCAACACATGAGGTGTCAGCTTCC 1028
   ||||| ||| ||||| : : : : : ||||| : : : : :
328 AAlaGlyGluITrThrCysLeuAlaGlyAsnSerIleGlyPheSerHisIT 344
   : : : : :
1029 CAGCGCTTCCTGACGCGTCGTCGA...GACCAAAACCGCGAGCGCAC 1075
   : : : : :
344 SerAlaITrPheValIThrValLeuProAlaGluGluLeuMetGluIThrA 361
   : : : : :
1076 CTGTGCTCTCTCTGTCGCGCCACTAGCTGCTGCGCGCTGTCATC 1125
   ||||| : : : : : ||| : : : ||||| : : :
361 sPGLAlaGlySerValITrAlaGlyValLeuSerITrGlyValAlaPhe 377
   : : : : :
1126 GGCATCCACGCGCGCTGTCTTCATCTGGGACACCTGCTCTGTGGCT 1175
   : : : : :
378 .....PheLeuPheIleLeuValAlaAlaValIThLe 389
   : : : : :
1176 TTGCGAGCGCCAGAGAGAGCGTCGACACCGCGCGCTGCGCCCTG 1224
   ||||| : : : : : ||| : : : ||||| : : :
389 uCysArgLeuArgSerProITrLysGlyLeuGlySerITrProIThrVal 406
   : : : : :
1225 .....CTTGGGACCGCGCGCGCGGAGCGCGCGCAC 1257
   : : : : :
406 IAsLysValSerArgPheProLeuLysArgGlnValSerLeuLysSerAsn 422
   : : : : :
1258 CGCAGCGGAGACAGACCTTCC.....TGTGTGGCGCGCTTACGCGC 1301
   ||| : : : : : ||| : : : ||||| : : :
423 SerSerMetAsnSerAsnIThrProLeuValArgIThAlaArgLysSerSe 439
   : : : : :
1302 TGGCCTGAGTGGGGCTGTGT.....GAGGAGACGATGCTCCGCGAG 1345
   : : : ||| ||| : : : : : ||||| : : :
439 rGlyGluGlyProValIThrValAlaAsnValSerITrLeuGlnITrProIla 456
   : : : : :
1346 CCCCCCAG 1353
   ||||| : : : : :
456 sPProLys 458
   : : : : :
seq_name: p1r1:TVH02F
seq_documentation_block:
  fibroblast growth factor receptor flg-2 precursor - human
  N:Contains: protein-tyrosine kinase (EC 2.7.1.112) flg-2
  C:Species: Homo sapiens (man)
  C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jun-2000
  C:Accession: A60350; S21843
  R:AviVl, A.; Zimmer, Y.; Yayon, A.; Yarden, Y.; GlyVol, D.
  Oncogene 6, 1089-1092, 1991
  A:Title: Flg-2, a new member of the family of fibroblast growth factor receptors.
  A:Reference number: A60350; MUID:91296390
  A:Accession: A60350
  A:Molecule type: mRNA
  A:Residues: 1-800 <AVI>
  A:Cross-references: EMBL:X58255; NID:931382; PIDN:CAA1209.1; PID:931383
  A:Experimental source: keratinocytes
  C:Comment: This may be a receptor for keratinocyte growth factor.
  C:Genetics:
  A:Gene: GDB:FGFR2; UMS; CPD1; KGF; Flg2
  A:Cross-references: GDB:127273; OMIM:176943
  A:Map position: 10q25.3-10q26
  C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; pr
  C:Keywords: Arp; autophosphorylation; duplication; glycoprotein; growth factor recept
  F:1-21/Domain: signal sequence #status predicted <Sig>
  F:22-800/Product: fibroblast growth factor receptor flg-2 #status predicted <MAT>
  F:22-369/Domain: extracellular #status predicted <EXT>
  F:131-137/Region: acidic
  F:262-335/Domain: immunoglobulin homology <IMM>
  F:370-390/Domain: transmembrane #status predicted <TM>

```







A:Reference number: A38576; MUID:91142118  
 A:Accession: A38576  
 A:Molecule type: mRNA  
 A:Residues: 1-806 <KE>  
 A:Cross-references: GB:M58051; NID:g182566; PIDN:AAA52450.1; PID:g182569  
 R:Thompson, L.M.; Plummer, S.; Schalling, M.; Altherr, M.R.; Gussell, J.F.; Housman, D.H  
 Genomics 11, 1133-1142, 1991  
 A>Title: A gene encoding a fibroblast growth factor receptor isolated from the Huntingt  
 A:Reference number: A55273; MUID:92147110  
 A:Accession: A55273  
 A>Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 76-394, 'V', 396-806 <THO>  
 A:Cross-references: GB:M64347; NID:g182564; PIDN:AAA58470.1; PID:g182565  
 A:Note: sequence extracted from NCBI backbone (NCBI:80296)  
 R:Partanen, J.; Maekela, T.P.; Allitalo, R.; Lehtesalo, H.; Allitalo, K.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 8913-8917, 1990  
 A>Title: Putative tyrosine kinases expressed in K-562 human leukemia cells.  
 A:Reference number: A38268; MUID:91062389  
 A:Accession: E38269  
 A:Molecule type: mRNA  
 A:Residues: 619-675 <PAR>  
 A:Cross-references: GB:M37782  
 R:Bellus, G.A.; Hefteron, T.W.; Ortiz de Luna, R.I.; Hecht, J.T.; Horton, W.A.; Machado,  
 Am. J. Hum. Genet. 56, 368-373, 1995  
 A>Title: Achondroplasia is defined by recurrent G380R mutations of FGFR3.  
 A:Reference number: I51880; MUID:95150025  
 A:Accession: I51880  
 A>Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 361-379, 'R', 381-415 <RES>  
 A:Cross-references: GB:S76733; NID:9914201; PIDN:AB33323.1; PID:9914202  
 A:Note: this sequence represents a mutant form associated with achondroplasia  
 C:Genetics:  
 A:Gene: GDB:FGFR3  
 A:Cross-references: GDB:127526; OMIM:100800; OMIM:134934  
 A:Map position: 4p16.3-4p16.3  
 C:Function:  
 A:Description: receptor for both acidic and basic fibroblast growth factors  
 C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote  
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;  
 F:1-22/Domain: signal sequence #status predicted <STG>  
 F:23-806/Product: fibroblast growth factor receptor 3 #status predicted <MAT>  
 F:23-375/Domain: extracellular #status predicted <EXT>  
 F:133-139/Region: acidic  
 F:268-341/Domain: immunoglobulin homology <IMM>  
 F:376-396/Domain: transmembrane #status predicted <TM>  
 F:397-806/Domain: intracellular #status predicted <INT>  
 F:470-755/Domain: protein kinase homology <KIN>  
 F:478-486/Region: protein kinase ATP-binding motif  
 F:61-109, 176-228, 275-339/Disulfide bonds: #status predicted  
 F:98, 225, 262, 294, 315, 328/Binding site: carbohydrate (Asp) (covalent) #status predicted  
 F:508, 525, 617/Active site: Lys, Glu, Asp #status predicted  
 F:622, 635/Binding site: magnesium (Asn, Asp) #status predicted  
 F:648/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicte

alignment\_scores:  
 Quality: 477.00 Length: 448  
 Ratio: 1.754 Gaps: 16  
 Percent Similarity: 60.714 Percent Identity: 31.027

alignment\_block:  
 US-09-598-042A-2 x TVHUF3 ..  
 Align seg 1/1 to: TVHUF3 from: 1 to: 806

```

76  CCGGGTCGACGATTCTGCGCAGGTGCGCCGCTGGCGCCGACACTATGCG 125
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
43  ProGlyGlnGln...GlnGlnLeuValPheGlySerGlyAspAlaValG1 58
126 GCTGCACTGCCCACTGGAGGGGACCCGCCG...CCGTCGACCATGT 169
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

58  uLeuSerCysProProProlGlyGlyGlyProMetGlyPro...ThrValT 74
170 GGACCAAGGATGGCCGACACCATCCACAGGCGGTGGAGCGCTTCGCGNG 219
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
74  rValValAspGlyThrGlyLeuValProSer...GlnrGlyValLeuVal 89
220 CTGCCGCGAGGGGCTGAAGGTGAAGCAGTGAAGCGGAGGATGCGGGCT 269
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
90  GlyProGlnArgLeuGlnValLeuAsnAlaSerHisGlyAspSerGlyVal 106
270 GTAAGTGTCC...AAGGCCACACAGCGCTTCGGCAGCTGACGCTCA 313
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
106  dYrSerCysArgGlnArgLeuThrGln.....ArgValLeuCysH 120
314 ACTACACCCCTGCTGCTGATGATGATGATGATGATGATGATGATGATG 363
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
120  IsPheSerValArgValThrAspAlaProSerSerGlyAspAspGlyAsp 136
364 GGGCCGACAGCTCTCTGTTGGGGGTCAAGAGACCCCGCAGCCAGCAGTG 413
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
137  GlyGlnAspGlyAlaGlnAspThrGlyValAspThrGlyValAlaProTyr 153
414 GGCACGACCGCGCTTCACACAGCCCTCCAAAGATGAGGCGCCGGGTGATG 463
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
153  P.....ThrArgProGlnArgMetAspLysLysLeuVal 165
464 CAGCGCCGCTGGGTAGCTCCGCTGAGCTCAAGTGCAGTGCAGCGGAC 513
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
165  lValAlProAlaAlaAsnThrValArgPheArgCysProAlaAlaGlyAsn 181
514 CCTGGCGCGGACATGACGTGATGAAGAGCAGCAGCGCTTGACG... 558
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
182  ProThrProSerIleSerTrpLeuLysAsnGlyArgGlnPheArgGlyG1 198
559  ...CGCCCAAGAGCGCGCTGAGCCGACGAAGAAGAAAGTACAGCAGG 604
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
198  uHisArgIleGlyGlyIleLysLeuArgHisGlnGlnTrpSerLeuValM 215
605 TGAAGAACCTGGCGCGGAGGACAGCGGCAATATACCTGCGCGCTGTG 654
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
215  ecGluSerValValProSerAspArgGlyAsnTyrThrCysValValGlu 231
655 AACCGCGCGGCGCCATCAACGCCACTTCAAGGTGATGATGATGATGATG 704
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
232  AsnLysPheGlySerIleArgGlnThrTyrThrLeuAspValLeuValu 248
705 GACCGCTTCAGCGCGGCTGCTGACAGCAGCAGCAGCAGCAGCAGCAGG 754
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
248  gSerProHisArgProIleLeuGlnAlaGlyLeuProAlaAsnGlnThrA 265
755 TGGACTTGGGGGGGACACAGTCTTCACAGTGAAGGTGGCGCAGCAGTG 804
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
265  lValLeuGlySerAspValGlnPheHisCysValValTyrSerAspAla 281
805 AAGCCGATGATCCAGTGGCTGAAGCCGCTGAGACTACGGCGCGCAGGCGG 854
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
282  GlnProHisIleGlnTrpLeuLysHisValGluVal..... 293
855 CCACAGCTCCACCATCATGATGGCGCGGCGCAGAAATTGTG...GTGCTGC 901
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
294  .AsnGlySerLysValGlyProAspGlyThrProTyrValThrValLeuL 310
310  ysrThrAlaGly.....AlaAsnThrThrAspLysGlnLeuGlnValLeu 324
952 CTGATCAACCGTCCCGCCAGAGCAGATGGCGGATGATGATGATGATGATG 1001
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
325  SerLeuHisAsnValThrPheGlnAspAlaGlyGlyTyrThrCysLeuAl 341
1002  CGCACAACCATGAGCTACAGTTCGACAGCGGCTTCACACGCTGCTGC 1051
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
341  agLysAsnSerIleGlyPheSerHisHisSerAlaTrpLeuValValLeuP 358

```

```

1052 CAGACCCAAAACCCGACAGGCCACCTGTGGCTCTGCTGCTCGGGCACT 1101
      ||
      ::
      ::
358 roLAgluglu.....GlueuValGluAlaAspIuAlaGly 370
      ::
      ::
1102 AGCTGCGCGTGGCCCGTGGTGCATGCCAGCCGGCGGTGC..... 1146
      |||
      |||
      |||
371 Ser.....ValTyrAlaGlyIleuSerTyrGlyValGlyLph 383
      ::
      ::
      ::
1147 .....TTCATCCTGGGCACCCCTGCTGCTGCTTTGCCAGGCCACA 1189
      ::
      ::
      ::
383 ePhelLeuPhelLeuValValAlaAlaValThrLeuCysArgLeuArg 400
      ::
      ::
      ::
1190 AGAAGCGGTGACCCCGCGCGCTGCCCT..... 1218
      ::
      ||
      ::
400 ePrProPolysLysGlyLeuGlySerProThrValHisLysIleSerArg 416
      ::
      ::
      ::
1219 ...CCCTGCGTGGGACACGCCGCCGCGGGAGCGGCCGCCAGCCGCG 1265
      |||||
      ::
      ::
417 PheProLeuLysArgGlnValSerLeuGluSerAsnAlaSerMetSer 433
      ::
      ::
      ::
1266 AGACAAAGACCTTCCTCGTGGCGCCCTCAGCGCTGGCGCCGTGGTGG 1315
      ::
      ||
      ::
433 rAsnThrProLeuValArgIleAlaArgLeuSerSerGlyGluGlyProt 450
      ::
      ::
      ::
1316 GGCTGTGT.....GAGGAGCATGGGTGCCGCGAGCCGCCAG 1353
      ||:::
      ::::
      ::::
450 hrLeuAlaAsnValSerGluLeuGluLeuProAlaAspPolys 464

```

```

seq.documentaln_block:
N:Contains: fibroblast growth factor receptor A2 precursor - African clawed frog
N:Contains: fibroblast growth factor receptor A2, short splice form; protein-tyrosine ki
C:Species: Xenopus laevis (African clawed frog)
C:Date: 08-Mar-1991 #sequence_revision 19-Jan-1999 #text_change 16-Jul-1999
C:Accession: A36477; C39752
R:Muscl, T.J.; Amaya, E.; Kirschner, M.W.
Proc. Natl. Acad. Sci. U.S.A. 87, 8365-8369, 1990
A:Title: Regulation of the fibroblast growth factor receptor in early Xenopus embryos.
A:Reference number: A36477; M0ID:91045998
A:Accession: A36477
A:Molecule type: mRNA
A:Residues: 1-812 <MUS>
A:Cross-references: GB:U24491; GB:M37201; NID:q857677; PID:AAA86868.1; PID:q857678
A:Note: 45-Trg, 49-Thr, 50-Val, and 64-Ser also found
R:Rieschel, R.; David, I.B.
Mol. Cell. Biol. 11, 2481-2489, 1991
A:Title: cDNA cloning and developmental expression of fibroblast growth factor receptors.
A:Reference number: A39752; M0ID:91203867
A:Accession: C39752
A:Molecule type: mRNA
A:Residues: 1-30,119-189,'G','191-418','L','420-636','R','638-788','V','790-812 <FRT>
A:Cross-references: GB:M63222; NID:q214899; PID:AAA49993.1; PID:q214900
C:Superfamily: Basic fibroblast growth factor receptor 1; Immunoglobulin homology; prote
C:Keywords: alternative splicing; ATP; autophosphorylation; duplication; glycoprotein; g
protein kinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:12-812/Product: fibroblast growth factor receptor A2, long splice form #status predict
F:12-372/Domain: extracellular #status predicted <EXT>
F:12-370,119-812/Product: fibroblast growth factor receptor A2, short splice form #status
F:447-102/Domain: Immunoglobulin homology <IM>
F:125-138/Region: acidic
F:167-128/Domain: Immunoglobulin homology <IM2>
F:167-228/Domain: Immunoglobulin homology <IM3>
F:266-339/Domain: transmembrane #status predicted <TM>
F:373-393/Domain: transmembrane #status predicted <TM>
F:394-812/Domain: intracellular #status predicted <INT>
F:470-755/Domain: protein kinase homology <KIN>
F:478-486/Region: protein kinase ATP-binding motif
F:554-100,174-226,275-337/Dissulfide bonds: #status predicted
F:76,116,133,177,223,236,260,292,313,326/Binding site: carbohydrate (Asn) (covalent) #st
F:508,525,617/Active site: Lys, Glu, Asp #status predicted
F:622,635/Binding site: magnesium (Asn, Asp) #status predicted

```

```

F:646/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status pred
alignment_scores:
    Quality: 469.00      Length: 437
    Ratio: 1.757        Gaps: 15
    Percent Similarity: 61.098      Percent Identity: 30.206
alignment_block:
    US-09-598-042A-2 x A36477 ..
Align seg 1/1 to: A36477 from: 1 to: 812

```

[illegible]





```

509 GGCACCTGGCCGACATCAGTGTGATGAGACGACCGCCCTGACG 558
183 LysnProthrSerThrMetArgTrpLeuLysAsnLysGluLeuLys 199
559 .....CGCCAGAGGCGCGTGGAGCCAGAGAGAGAGAGTGGACCT 599
200 GlnGlnHisArgIleGlyLeuLysValArgAsnGlnHisTrpSerLe 216
600 GAGCCTGAAGAACCTCGCCCGGAGACAGCGGCAATACACCTGCCGG 649
216 LileMetGlnSerValValProSerAspLysGlnAsnTrpLysLeuV 233
650 TGTGCAACCGCGCGCGCATCAGACGACGACGACGAGGATGATGATC 699
233 ALGlnAsnGluTrpGlySerIleAsnHisThrTrpHisLeuAspVal 249
700 CAGCGGACCGCTTCAAGCCGCTGTACAGGACGACCGCCGCTGAAC 749
250 GluArgSerProHisArgProIleGlnIleGlnAlaGlyLeuProAla 266
750 GACGGTGGACTTGGGGGAGACGACGCTTCCAGTGCAGAGTGCAGCG 799
266 AserThrValAlaGlyLysPvalGlnPvalCysLysValTrpSer 283
800 ACCTGAGCGCGTGTATCAGTGTGATGAGCGCGCTGAG ..... 837
283 sPrlaGlnProHisIleGlnTrpIleGlnHisValGlnLysAsnGly 299
838 ...TACGGCGCGAGGCGGCGCACAACTCCACATGATGTGGCGGCA 864
300 LysTrpGlyProAspGly ..... 305
885 GAGTTTGTGTGTGTCGCCGAGGTGACGTTGTGCGGCGCGAGGCT 934
306 .....LeuProTrpLeuLysValLeuLysAlaAlaGlyVal 318
935 CCTACCTCATTAAG .....CTGTCTATCACCCTGTCGCCGCGCAG 972
318 snThrTrpAspLysGlnIleGluValLeuTrpIleArgAsnValThrPhe 334
973 GACGATGCGGGCATGTACATCTGCTTGGCGGACACACCTGGGCTAC 1022
335 GluAspAlaGlyLeuTrpThrCysLeuAlaGlyAsnSerIleGlyIle 351
1023 CTTCCGAGGCGCTTCCCTCAGCGTGCAGCCGCAACCCGACGAGGC 1072
351 rPheHisSerAlaTrpLeuThrValLeuProAlaPro ..... 363
1073 CACCTGTGGCTCTCTGCTCGGCCACTAGCCTGCGGCGCGGTGTC 1122
364 .....ValArgGluLysGlnIleThrAlaSerProAspTrpLeuGlu 377
1123 ATGGGCACTCCAGCGCGCTGTCTTATCTTGGGACACCTGCTCGTGTG 1172
378 IleAlaIleTrpCysIleGlyValPheLeuIleAlaCysMetValAla 394
1173 GCTTGGCAGGCGGAGAGAGCGGTGACCCCGCGGCGCT 1212
394 rValIlePheCysArgMetLysThrTrpLysPro 407

seq_name: p1r2:B35963

seq documentation block:
protein-tyrosine kinase (EC 2.7.1.112) cek3 precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 16-Jul-1999
C:Accession: B35963
R:Pasquale, E.B.
Proc. Natl. Acad. Sci. U.S.A. 87, 5812-5816, 1990
A:Title: A distinctive family of embryonic protein-tyrosine kinase receptors.
A:Reference number: A35963; MUID:90332672
A:Accession: B35963
A:Status: preliminary

```

```

A:Molecule type: mRNA
A:Residues: 1-823 <PAS>
A:Cross references: GB:M35196; NID:g211444; PIDN:AAA4865.1; PID:g211445
C:Genetics:
A:Gene: cek3
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; pr
C:Keywords: Atp; autophosphorylation; duplication; glycoprotein; growth factor recept
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-823/Product: protein-tyrosine kinase cek3 #status predicted <MNT>
F:23-379/Domain: extracellular #status predicted <EXT>
F:135-141/Region: acidic
F:174-235/Domain: immunoglobulin homology <IMM>
F:380-400/Domain: transmembrane #status predicted <TM>
F:401-823/Domain: intracellular #status predicted <INT>
F:481-766/Domain: protein kinase homology <KIN>
F:489-497/Region: protein kinase ATP-binding motif
F:65-110,181-233,280-344/Disulfide bonds: #status predicted
F:230,267,299,320,333/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:519,536,628/Active site: Lys, Glu, Asp #status predicted
F:633,646/Binding site: magnesium (Asn, Asp) #status predicted
F:659/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi

alignment_scores:
Quality: 463.00 Length: 463
Ratio: 1.802 Gaps: 13
Percent Similarity: 55.508 Percent identity: 27.646

alignment_block:
US-09-598-042a-2 x B35963 ..

Align seg 1/1 to: B35963 from: 1 to: 823

112 GGCACGACATGAGCGGTGACATGTCAGTGCAGGAGGAGCGCCGCGCT 161
58 GlyLeuProLeuGlnLeuArgCysGlnLeuLys .....AspAlaValMet 72
162 GACCATGTGAGCAGCAAGATGCG .....CGCACCA 190
72 tileSerTrpThrLysAspGlyValProLeuGlyProAspAsnArgThrV 89
191 TCACACGCGGCTGAGACCGCTTCCGCTGCTGCCGAGGCGCTGAAGTG 240
89 al .....IleIleGlyLeuTrpLeuGlnIle 97
241 AACGAGTGTGAGCGGAGATGCGCGGCTGTACGTGTGCAAGCCACCA 290
98 LysAspAlaSerProArgAspSerGlyLeuTrpAlaCysThrAlaIleAr 114
114 gThrLeuAspSerAspThrLeuTrpPheIleValAsnValThrAspAla 131
341 TTTAGCCAGGAGAGAGAGACCTTGGGCGCGACAGCTTCCTGGGCTCA 390
131 euserSerGlyAspAsp .....GluAspAspAsnAspGlySer 143
391 GAGGAC .....CCGCCAGCCAGCATGAGGAGCGACGCGCGCTTCACACA 434
144 GluAspPheValAlaAsnAspSerAsnGlnMetArgAlaProTrpTrpThr 160
435 GCCCTCCAGATGAGGCGCGGTGATGCGACGCGCGGTGGATAGCTCG 484
160 snThrAspLysMetGlnLysArgLeuHisAlaValProAlaAlaAsnThrV 177
485 TCGCGCTCAAGTGTGCGCGAGGCGGACCTTGGCGCGCACATCAGCTG 534
177 alLysPheArgCysProAlaMetGlyAsnProThrProThrMetLysGly 193
535 ATGAGAGAGCAGCAGGCGCTTGACG .....CGCCAGAGGCGCGCTGA 575
194 LeuLysAsnGlyLysGluPheLysGlnGlnHisArgIleGlyLysTrp 210

```



Quality: 462.00 Length: 387  
Ratio: 1.958 Gaps: 10  
Percent Similarity: 60.982 Percent Identity: 31.008

alignment\_block:

us-09-598-042a-2 x TVHUF2 ..

Align seg 1/1 to: TVHUF2 from: 1 to: 821

```

100 GTGGCCGGCTGGGCGGCGGCTGATGCGGCTGACGTCAGTGGCCAGTGGAGGGGGA 149
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51 ValAlaIaIaProGluSerLeuGluValAlaGcysLeuLys..... 65
150 CCGCGCGCGCGCTGACATGTGACCAAGATGGCGCGGCGGCGGCTGACAGGC 199
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
66 .AspAlaIaValIleSerThrThrLysAspGly.....ValHisLeuG 80
200 GCTGGAGCGCTTCCGCGTCTG...CCGCGAGGCGCTGAAGTGAAGCAG 246
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
80 LysProAsnAsnArgThrValLeuIleGlyGluLeuGlnIleLysGly 96
247 GTGGAGCGGAGATGCGGCGGCTGATGCTGACGTCAGGCGGCGGCTGACGCTT 296
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
97 AlaThrProArgAspSerGlyLeuTyAlaCysThrAlaSerArgThrVa 113
297 CCGCAGCGCTGAGCTCACTACACCTCTGCTGATGATGATGATGATGATGATG 346
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
113 LaspSerGluThrThrPyrThrMetValAsnValThrAspAlaIleSer 130
347 CAGGGAAGAGAGCTGGGCGGCGGCGGCTGCTGGGCGGCTGACAGAGAC 396
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
130 eGlyAspAsp.....GluAspAspThrAspGlyAlaGluAsp 142
397 CCGCGCGCGGCGGCTGACGTCAGGCGGCGGCTGACAGAGCGGCTGACAGCGC 437
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
143 PheValSerGluAsnSerAsnAsnLysAlaAlaProTyrThrAsnThr 159
438 CTCGAGATGAGCGCGGCGGCTGATGCGGCGGCGGCGGCTGATGCTGCTG 487
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
159 rGluLysMetGluLysArgLeuHisAlaValProAlaAlaAsnThrVal 176
488 GCGTCAAGTCTGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
176 yspPheArgCysProAlaGlyLysProMetProThrMetArgTyrLeu 192
538 AAGGAGCAGCAGCGCTGAGC.....CGCCGAGGCGGCGGCTGAGGCC 578
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
193 LysAsnGlyLysGluPheLysGlnGlnHisArgIleGlyLysTyrLysVa 209
579 CAGGAGAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 628
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
209 LArgAsnGlnHisThrSerLeuIleMetGluSerValValProSerAsp 226
629 GCGGCAAAATGACCTGCGGCTGCAAGCGGCGGCGGCGGCGGCGGCGGCGG 678
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
226 yspGlyAsnTyrThrCysValAlaGluLysGlnIleGlySerIleAsnHis 242
679 ACCTAGCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 728
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
243 ThrTyrHisLeuAspValAlaGluArgSerProHisArgProIleLeuG 259
729 AGGCGACGACCGCTGAGACGAGCGGCTGATGCGGCGGCGGCGGCGGCGGCT 778
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
259 nAlaGlyLeuProAlaAsnAlaSerThrValAlaGlyGlyAspValGlu 276
779 TCACATGCAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 828
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
276 heValCysLysValTyrSerAspAlaIleProHisIleGlnTyrIleLys 292
829 CCGGTGGAG.....TACGGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGG 863
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
293 HisValGluLysAsnGlySerLysTyrGlyProAspGly..... 305

```

```

864 CACCATGATGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 913
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
306 .....LeuProTyrLeuLysVal 311
914 TGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 951
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
311 alLeuLysAlaAlaGlyValAsnThrThrAspLysGluIleGluValLeu 327
952 CTCATGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1001
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
328 TyrIleArgAsnValThrPheGluAspAlaGlyGlyTyrThrCysLeuAl 344
344 acGlyAsnSerIleGlyIleSerPheHisSerAlaThrPleuThrValLeu 361
1052 CAGACCCAAACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1101
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
361 roAlaPro.....GlyArgGluLysGluIleThr 370
1102 AGCGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 1151
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
371 AlaSerProAspTyrLeuGluIleAlaIleTyrCysIleGlyValPheLe 387
1152 CCGTGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1201
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
387 uIleAlaCysMetValValThrValIleLeuCysArgMetLysAsnThrT 404
1202 CCGCGCGGCGCT 1212
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
404 hLysLysPro 407

```

seq\_name: p1r1.TVHUF4

```

seq_documentation_block:
Fibroblast growth factor receptor 4 precursor - human
N:Alternate names: protein-tyrosine kinase tkf
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text change 16-Jul-1999
C:Accession: S15345; A46615; A41598; D38269
R:Partanen, J.; Maekela, T.P.; Betola, E.; Korhonen, J.; Hirvonen, H.; Claesson-Wells
EMBO J. 10, 1347-1354, 1991
A:Title: FGFR-4, a novel acidic fibroblast growth factor receptor with a distinct exp
A:Reference number: S15345; MUID:91224085
A:Accession: S15345
A:Molecule type: mRNA
A:Residues: 1-802 <PAB>
A:Cross-references: EMBL:X57205; NID:g31371; PIDN:CAA40490.1; PID:g31372
A>Note: binds acidic but not basic fibroblast growth factor with high affinity
R:Ron, D.; Reich, R.; Chedid, M.; Lengel, C.; Cohen, O.E.; Chan, A.M.; Neufeld, G.; M
J. Biol. Chem. 268, 5388-5394, 1993
A:Title: Fibroblast growth factor receptor 4 is a high affinity receptor for both ac1
A:Reference number: A46615; MUID:93194827
A:Accession: A46615
A>Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-296, 'D', 298-802 <RON>
A:Experimental source: mammary epithelial cell line B5/589
A>Note: sequence extracted from NCBI backbone (NCBI:127650)
A>Note: binds acidic and basic fibroblast growth factor with high affinity
R:Hottrich, U.; Braeuninger, A.; Streubardt, K.; Ruesamen-Waigmann, H.
Proc. Natl. Acad. Sci. U.S.A. 88, 10411-10415, 1991
A:Title: Two additional protein-tyrosine kinases expressed in human lung: fourth memb
A:Reference number: S19025; MUID:92073297
A:Accession: A41598
A>Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 399-534, 'W', 536-799, 'SG', 800-802 <HOL>
A:Experimental source: lung
R:Partanen, J.; Maekela, T.P.; Allitalo, R.; Lehtsaialho, H.; Allitalo, K.
Proc. Natl. Acad. Sci. U.S.A. 87, 8913-8917, 1990
A:Title: Putative tyrosine kinases expressed in K-562 human leukemia cells.

```



A:Reference number: A38268; MUID:91062389  
A:Accession: D38269  
A:Molecule type: mRNA  
A:Residues: 614-670 <PA2>  
A:Cross-references: GB:M3781  
A:Experimental source: K-562 leukemia cell line  
C:Genetics:  
A:Gene: GDB:FCGR4  
A:Cross-references: GDB:127929; OMIM:134935  
A:Map position: 5q33.2-5qter  
C:Function:  
A:Description: receptor mediating effects of fibroblast growth factor  
A:Note: expressed in normal lung; expressed in some carcinomas  
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; proteoglycan  
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-802/Product: fibroblast growth factor receptor 4 #status predicted <MAT>  
F:25-365/Domain: extracellular #status predicted <EXT>  
F:50-103/Domain: immunoglobulin homology <IM1>  
F:165-226/Domain: immunoglobulin homology <IM2>  
F:264-335/Domain: immunoglobulin homology <IM3>  
F:370-390/Domain: transmembrane #status predicted <TMN>  
F:391-802/Domain: intracellular #status predicted <INT>  
F:465-750/Domain: protein kinase homology <KIN>  
F:473-481/Region: protein kinase ATP-binding motif  
F:57-101,112-224,271-333/Disulfide bonds: #status predicted  
F:112,268,290,311,322/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:503,520,612/Active site: Lys, Glu, Asp #status predicted  
F:617,650/Binding site: magnesium (Asn, Asp) #status predicted  
F:643/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

alignment_scores:		
Quality:	459.50	Length: 409
Ratio:	1.947	Gaps: 15
Percent Similarity:	57.702	Percent Identity: 31.785

alignment\_block:

US-09-598-042A-2 x TVHUE4

Align seg 1/1 to: TVHUF4 From: 1 to: 802

[illegible][illegible]

C:Species: Gallus gallus (chicken)  
 C:Date: 09-Nov-1990 #sequence\_revision 09-Nov-1990 #text\_change 16-Jul-1999  
 C:Accession: A35963  
 R:Pasquale, E.B.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 5812-5816, 1990  
 A:Title: A distinctive family of embryonic protein-tyrosine kinase receptors.  
 A:Reference number: A35963; MUID:90332672  
 A:Accession: A35963  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-806 <PAS>  
 A:Cross-references: GB:M55195; NID:9211442; PID:AAA48664.1; PID:9211443  
 C:GeneInfo:  
 A:Gene: cek2  
 C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote  
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;  
 F:1-73/Domain: signal sequence #status predicted <SIG>  
 F:24-806/Product: protein-tyrosine kinase cek2 #status predicted <MAT>  
 F:131-138/Region: extracellular #status predicted <EXT>  
 F:131-138/Region: acidic  
 F:262-335/Domain: immunoglobulin homology <IMM>  
 F:369-389/Domain: transmembrane #status predicted <TM>  
 F:390-806/Domain: intracellular #status predicted <INT>  
 F:464-749/Domain: protein kinase ATP-binding motif  
 F:472-480/Region: protein kinase ATP-binding motif  
 F:61-107,170-222,269-333/Disulfide bonds: #status predicted  
 F:96,219,256,288,309,322/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:502,519,611/Active site: Lys, Glu, Asp #status predicted  
 F:616,629/Binding site: magnesium (Asn, Asp) #status predicted  
 F:642/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

alignment\_scores:  
 Quality: 456.50 length: 440  
 Ratio: 1.723 Gaps: 13  
 Percent Similarity: 60.227 Percent Identity: 30.227

alignment\_block:  
 US-09-598-042a-2 x A35963 ..

Align seg 1/1 to: A35963 from: 1 to: 806

```

112 GGGCGACATATGCGGCTGCAGTCCAGTGGAGGGAGCCGCCGCGT 161
    ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:
54 GLYSPHRLIEGLIULSERCYASNTHGLINSERSEVALSERVA 70
162 GACCATGTGACCAAGAGTGGCGGACACATCCAGGGCGCTGAGCGCT 211
    : |||:|||||:|||||:|||||:|||||:|||||:|||||:
70 L...PHEITRPHLYSPGLYLLIEGLYIEALPRTSER...ASNRTGT 85
212 TCCGCTGTGTCGCCGAGGGCTGAAGGTGAAGCAGGTGAGCGGAGAT 261
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:
85 HRLISLIEGLYGLNLSLEULEYSLIELLEASNVALSERLYRASPASP 101
262 GCCGCGGTGATGCTGTGCAAGGCCACCAAGCGCTTGCGGACGCTGAGCT 311
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:
102 SERGLYLEUTYSERCYSLSPROARGHISSEKASGLIUVALLLEUGLY.. 117
312 CAACCTACACCTCTGCTGCTGATGATGATATACCCAGGAGAGAGAGCC 361
118 .ASNPETHIVALRGVALTHRASP..... 125
362 TGGGGCCGACAGCTCTCTGGGGGTCAAGAGAGCCCGCCGACGACAG 411
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:
126 ..SERPROSERSEGLYASPAAPGLUASPASPASPASPASPGLUSERGLUASP 141
412 TGGGACGACGCGGCTTACACAGCCCTCCAAAGATGAGCGCGGCTGAT 461
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:
142 THGGLYVALPRTOPHETPRTHRARGPROASPLYSMETGLULYSLEULE 158
462 CGACGCGCGGTGGTAGCTCGGAGGGGCTCAAGTGGTGGCCAGCGGGC 511
158 UALAVAIPIOTALALASNTHRVALNRPHNARGCYSPROALAGLYGLYA 175

```

```

512 ACCSTGGSCCGACATACAGTGGATGAAGAGCAGCAGCGCTTGAGC... 558
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:
175 SNPRTHRPRTHRLIEYTRPRLEULYSASNGLYSGULPRHLYSGLY 191
559 .....CGCCAGAGCGCGCTGAGCCGAGAGACAGACAGTGGACAGT 602
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
192 GLUNHISATGLIEGLYILEYLSLEUARGHISGLNGLNTPRSEKLEUVA 208
603 CSTGAGAACCTGGCGCGGAGGAGCAGCGGCAATACCTGCGCGGTGT 652
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:
208 IMETGLUSERVALVALPRTSERASPRARGLYASNLYTRHLYCYVALVALG 225
653 CGAAGCGCGGCGGCGGACATACAGCCACCTACAGGTGATGATTCAG 702
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:
225 LUNSLNYSYGLYASNLIENRHNISLTHLYGCLINLEUASPARVALLLEUGLU 241
703 CGGACCGCTTCCAGCGCGCTGCTACAGGACGACCGCCGCTGAACAGAC 752
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
242 ARGSEPRONHISARPRDILEUGINALAGLYLEUPROALASNGLNTH 258
753 GRTGAGACTTGGGGGAGACACAGTCTCCAGTGCAGAGTGGCAGGAGC 802
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
258 RVALVALVALGLYSEKASNAVALIUPHENISCYLSVALYTRSERASPA 275
803 TGAAGCGGATGATCAGATGCTGACAGCGCGCTGAGC..... 837
275 IALINPRONHISLIEGLNTPRLEULYHNSIVALDLVALASNGLYSELYS 291
838 TACGCGCGGAGGCGCGGACCAACCTCCACCATGATGTGGCGGCCAGAA 867
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
292 TYRGLYPROASPLIYTHRPROLYVALTHR..... 301
888 GTTGTGCTGCTGCCACGAGGTGACGNTGTGCGGCGCGAGCGCTCT 937
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:
302 .....VALLEULYSTHRLAAGLYVAL.....ASNTHHASPRLYS 314
938 ACCTCAATAGACTCTCATACACCGCTGCCGCGGACGATGCGGCGCAT 987
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
314 IULUUGLIULEULYTRLEUARGASNVALTRHPEGIULASPARIALAGLYLU 330
988 TACATGCTGCTTGGCGGCAACACCATGGGCTACAGCTTCGCGAGCGCTT 1037
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
331 TYRTHCYSLUENALIALYSNSETLIEGLYRPHSETHISHSERLATTR 347
1038 CTTACCGCTGCTGCGAAGCCAAACCGGACGAGCGGACCTGTGCGCTCT 1087
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:
347 PLEUTHRVALLEUPROALAGLIULUULEUMETGLUMETASPARSERGLYS 364
1088 CGTCTCGGCGACTACCTGCTGCCGCGCGCTGTCATCGGCATCCAGACC 1137
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
364 ERYALYTRALAGLYLEULEUSERYGLYTHRGLYU..... 376
1138 GGGCGTGTCTCATCTGCGGACACCTGCTGCTGCTTGGCAGGCGCA 1187
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:
377 ...VALLEUHELLEUVALLEUVALILEVALILELEYSARG...ME 391
1188 GAGAGAGCGG.....TCACCGCGCGCGCTGCGC 1216
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:
391 TLYSMETPRASNLNLYSLALAMETASNTHRTHRTHRVALGHLNLYVALS 408
1217 CTGCGCTGCTGGGACACCGCGCGCGGAGCGGCGGACGCGCAGC... 1263
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:
408 ERLYSPHEPRLEULYSARGGLNGLNVALSERLEUULSERASNSER 424
1264 .....GGAGACAAGACCTTCCCTGTTGGCGCGCTGACGCGCTGGCCC 1307
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:
425 METASNSERASNTHRPROLEUVALARGILETTHRYLEUSERSEKAS 441
1308 TGGTGTGGGCGCTGTGAGAGAGCATGGGTCTCGGAGCGCGCCAGCACT 1357
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
441 PGLYPRMETLEU..ALASNVALSERGLIULUENGLIULUENPROPRDAPPRO 457

```

1358 TACTGGGCCAGGCCAG 1375  
 458 LYSTPGLULEuAlaArg 463  
 seq\_name: p1r2:B54846  
 seq\_documentation\_block:  
 fibroblast growth factor receptor b precursor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:date: 21-Jul-1995 #sequence\_revision 28-Jul-1995 #text\_change 16-Jul-1995  
 C:Accession: B54846  
 R:Tagaki, Y.; Shrivastav, S.; Miki, T.; Sakaguchi, K.  
 J. Biol. Chem. 269, 23743-23749, 1994  
 A:title: Molecular cloning and expression of the acidic fibroblast growth factor receptor  
 A:Reference number: A54846; MUID:94375484  
 A:Accession: B54846  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-822 <TRK>  
 C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote  
 C:Keywords: ATP; growth factor receptor  
 F:172-233/Domain: immunoglobulin homology <IMM>  
 F:480-765/Domain: protein kinase homology <KIN>  
 F:488-496/Region: protein kinase ATP-binding motif

alignment\_scores:  
 Quality: 455.50 Length: 378  
 Ratio: 1.947 Gaps: 9  
 Percent Similarity: 61.905 Percent Identity: 31.217

alignment\_block:  
 US-09-598-042A-2 x B54846 ..  
 Align seg 1/1 to: B54846 from: 1 to: 822

```

112 GCGCGCATATGCGGCTGCTGAGTGGAGGGGAGCCGCCGCT 161
    |||.....|.....|.....|.....|.....|.....|
55 GlySerLeuLeuLeuLeuArgCysMetLeu.....AspAlaAlaVal 69
162 GACCATGTGGACCAAGATGGCGCCAGCATGACAGCGGTGGAGCGCT 211
    |||.....|.....|.....|.....|.....|.....|
69 LLeSerTrpThrLysAspGly.....ValHisLeuGlyProAsnAsn 84
212 TCCCGCTGCTG...CCGCGAGGCTGAGTGAAGCAGGTGAGCGGAG 258
    |||.....|.....|.....|.....|.....|.....|
84 rGlnValLeuIleGlyLeuTrpLeuGlnIleLysGlyAlaThrProArg 100
259 GATGCGCGGCTGTACGTGTGCAAGGCCACCAAGCGCTTGGCAGCTGAG 308
    |||.....|.....|.....|.....|.....|.....|
101 AspSerGlyLeuTrpLacCysAlaAlaAlaValThrValAspSerGly 117
309 CGTCAACTACACCCGCTGCTGCTGATGATGATGAGCCAGGAGAGA 358
    |||.....|.....|.....|.....|.....|.....|
117 rLeuTrpThrMetValAsnValThrAspAlaIleSerSerGlyAspArg 134
359 GCCTGGGGCCGACAGCTCTCTGGGGGTCAAGAGACCCGCGAGCAG 408
    |||.....|.....|.....|.....|.....|.....|
134 LuAspAspThrAspThrSerGlnAspPheValSerGlnAsnArgSerAsn 150
409 CAGTGGGACAGCCGCGCTTACACAGCCCTCCAAAGATGAGCGCGGCT 458
    |||.....|.....|.....|.....|.....|.....|
151 GlnArgAla...ProTrpTrpThrAsnThrGlnLysMetGlnLysArg 166
459 GATGCGAGCGCCCTGGGTAGCTCCGCTGAGTCAAGTGGCGTGGCAGC 508
    |||.....|.....|.....|.....|.....|.....|
166 uHisAlaValProAlaAlaAsnThrValLysPheArgCysProAlaGly 183
509 GGCACCTCGGCGGACATGACGTGATGAGACGACGAGCGCTTGACG 558
    |||.....|.....|.....|.....|.....|.....|
183 LysAsnProThrProThrMetArgTrpLeuLysAsnGlyLysGlnPheLys 199

```

```

559 .....CGCCAGAGGCGCGTGAAGCCAGAGAGAGAGAGTGAACACT 599
200 GlnGlnHisArgIleGlyLeuTrpLysValArgAsnGlnHisTrpSer 216
600 GAGCCTGAAGAACTGTCGCGCGAGAGACAGCGCAATATACACTGCCG 649
    |||.....|.....|.....|.....|.....|.....|
216 uIleMetGlnSerValValProSerAspLysGlnLysTrpThrCysLeu 233
650 TGTGACACCGGCGGCGCGCATCAAGCCGCACTACAGGTGATGTGATC 699
    |||.....|.....|.....|.....|.....|.....|
223 AlGlnAsnGlnLysGlySerIleAsnHisThrLysHisLeuAspVal 249
700 CAGCGGACCCGCTTCAAGCGCGCTGCTCAAGGCGAGCGACCCGCTA 749
    |||.....|.....|.....|.....|.....|.....|
250 GlnArgSerProHisArgProIleLeuGlnAlaGlyLeuProAlaAsn 266
750 GACGCTGACATTCGGGGGACACGCTCTTCAGTCAAGTGGCGACG 799
    |||.....|.....|.....|.....|.....|.....|
266 AserThrValValGlyGlnAspValGlnPheValLysLysValTrpSer 283
800 ACGTGAAGCGGCTGATCCAGTGGTGAAGCGGCTGAG..... 837
    |||.....|.....|.....|.....|.....|.....|
283 sPAlaGlnProHisIleGlnTrpIleLysHisValGlnLysAsnGlySer 299
838 ...TACGCGCGCGAGCGCGCCACAACTCCACATCGATGTGGCGGCA 884
    |||.....|.....|.....|.....|.....|.....|
300 LysTrpGlyProAspGly..... 305
885 GAAGTTGTGGTGTGCCACAGGCTGACGTG.....TGTGCGCGCGCG 928
    |||.....|.....|.....|.....|.....|.....|
306 .....LeuProTrpLeuLysValLeuLysHisSerGlyIleAla 318
929 ACGGCTCTACCTCAATAAGCTGTGTATCACCCTGCGCGCGAGACGAT 978
    |||.....|.....|.....|.....|.....|.....|
318 snSerSerAsnAlaGlnValLeuAlaLeuPheAsnValThrGlnMetAsp 334
979 GCGGCGATGTATATCTGCTTGGCGCGCAACACATAGGCTAAGCTTCC 1028
    |||.....|.....|.....|.....|.....|.....|
335 AlaGlyGlnTrpIleCysLysValSerAsnTrpTrpIleGlnAlaAsn 351
1029 CAGGCGCTTCTCTACCGCTGCTGACACCCAAACCGCGGCGCACCTG 1078
    |||.....|.....|.....|.....|.....|.....|
351 nSerAlaTrpLeuThrValLeuProLysGlnAlaPro..... 364
1079 TGGCCTCTGCTCTGCGGACATGACCTGCGTGGCGCGGTGTCATCGC 1128
    |||.....|.....|.....|.....|.....|.....|
365 ...ValArgGlnLysGlnIleThrAlaSerProAspTrpLeuGlnIleAla 380
1129 ATCCAGCCGCGGCTGTCTTCACTGCGGACACCTGCTCTGTGGCTTGG 1178
    |||.....|.....|.....|.....|.....|.....|
381 IleTrpCysIleGlyValPheLeuIleAlaCysMetValValThrVal 397
1179 CCAGGCGCCAGAGAGCGGTGCACCCCGCGGCT 1212
397 ePheCysArgMetLysTrpThrThrLysLysPro 408

```

seq\_name: p1r2:JC4583  
 seq\_documentation\_block:  
 fibroblast growth factor receptor 4B precursor - African clawed frog  
 C:Species: Xenopus laevis (African clawed frog)  
 C:date: 10-Apr-1996 #sequence\_revision 24-May-1996 #text\_change 16-Jul-1999  
 C:Accession: JC4583  
 R:Riou, J.F.; Clavilier, L.; Boucancut, J.C.  
 Biochem. Biophys. Res. Commun. 218, 198-204, 1996  
 A:title: Early regionalized expression of a novel Xenopus fibroblast growth factor re  
 A:Reference number: JC4583; MUID:96136300  
 A:Accession: JC4583  
 A:Molecule type: mRNA  
 A:Residues: 1-829 <RIO>  
 A:Cross-references: EMBL:X89807  
 A:Experimental source: embryo  
 C:Comment: This receptor is a tyrosine kinase transmembrane protein and plays an impo

C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote

C:Keywords: ATP; embryo; fibroblast; growth factor; receptor; transmembrane protein

F:1-26/Domain: signal sequence #status predicted <SIG>

F:27-829/Product: fibroblast growth factor receptor 4B #status predicted <MAT>

F:60-117/Domain: immunoglobulin homology <IM1>

F:136-148/Region: acidic

F:187-251/Domain: immunoglobulin homology <IM2>

F:287-362/Domain: immunoglobulin homology <IM3>

F:393-413/Domain: transmembrane #status predicted <TM>

F:488-773/Domain: protein kinase homology <KIN>

F:496-504/Region: protein kinase ATP-binding motif

alignment\_scores:

Quality:	456.00	Length:	405
Ratio:	1.861	Gaps:	12
Percent Similarity:	60.494	Percent Identity:	29.630

alignment\_block:

us-09-598-042a-2 x JC4583

Align seg 1/1 to: JC4583 from: 1 to: 829

```

68 TGTGCTCCCGGTCGACGATTCGTCGCGAGGTGCGCGCTG.....111
||||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
31 CysArgProArgLeuSerGluAspGluAlaAsnTrpLysGlnThrI 47
112 .....GGCCGCACATATGC 124
47 urTrSerGluValGlnGluHisLeuLeuAspProGlyAsnAlaLeu 64
125 GGTGCGATGTCGCCAGTGTAGGGGACCGCGCGGTGACCATGTGGACC 174
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
64 rGluPheCys.....AspThrAsnGlnSerSerSerIleAsnTrpTyr 78
175 AAGATGCGCCGACCATATCAGACGCGGTGAGCGCTTCGCGGTGCC 224
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
79 ArgGlnGlnIuArgLeuLeuSerGly...GlyLysIleArgMetValG 94
225 GCAAGGCTCAGGTGAAGCAGGTGAGCGGAGAGATGCGCGGTGACG 274
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
94 yTrHisValLeuGlnValSerAspValThrTrpGluAspSerGlyLeu 111
275 TGTGCGAGGCGCACCAAGCGCTTCGCGAGCGTCACTACACCTGC 324
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
111 Iecys...ValValArgGlyThrGlyLysIleLeuArgArgPheSerIle 126
325 GTGCTGTGATGACATTAGCCAGG.....AAGAGACCTGGG 365
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
127 SerValValAspSerIleuAlaSerGlyAspGlnGluGlnGluAs 143
366 GCCCGAAGCTCCTCTGGGGGTCAAGAGACCCGCCAGCAG..... 408
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
143 pAspAspAspGluAspGlnArgArgGluAspThrThrAlaAspIle 160
409 .....CAGTGGCAGCAGCGCGCTTCACACAGCCCTTCAAG 444
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
160 IuGlnProValTyrPhePheGlnAlaLeuArgTrpTrpGlnProSer 176
445 ATGAGGCGCGGGGTGATGCGACGCGCGGTGATGCTCGCGCTCAA 494
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
177 MetAspLysLysLeuHisAlaValAlaProAlaGlyAsnThrVal 193
495 GTGGCTGGCCAGCGGACCTCGCGCCGACATACCTGATGAAGACG 544
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
193 rCysProAlaIleGlySerProLeuProThrIleArgTrpSerLeu 210
545 ACCAGGCTTGACCGCCAGAG.....GCCGCGACCGCCAGG 582
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
210 snGlyArgGlnPheArgGlyLysIleHisArgIleGlyIleGlnIle 226
583 AAGAAGAGTGGACATGAGCCTGAGAGACCTGGCGCGAGACGACGG 632
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||

```

```

227 HisGlnHisTrpSerLeuValMetGluSerValValProSerAspArg 243
633 CAATATCACCTGCCGCTGTCGACCGCGCGGCGCCATCAACGCACT 682
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
243 yAsnTrpTrpCysValValGlnAsnArgValGlySerLeuThrTyr 260
683 ACAAGGTGATGATCCAGCGGACCGCTTCCAAAGCCCGTGCAGAGC 732
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
260 yTrHeuAspValLeuGlnArgSerSerHisArgProIleLeuGlnAl 276
733 ACACACCCCGTGAACAGACGCTGAGCTTGGGGGACACAGCTCTTCA 782
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
277 GlyLeuProAlaAsnThrThrAlaArgValGlySerAspValGln 293
783 GTGCAAGCTGCGGACGAGCTGAAGCGGTGATCCAGTGGCTGAACG 832
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
293 rCysLysValTyrSerAspAlaGlnProHisIleGlnTrpLeuLys 310
833 TGGAGTACGCGCGGAGCGCGCGCCACACACTCCACATCATGTGG 882
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
310 IeGlu.....ValAsnGlySerHis.....PheGlyPro 319
883 CAGAGATT.....GTGCTGCTGCCCGACGGGTGACGTGTGCGG 923
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
320 AspAspPheProTyrValGlnValLeuLysThrAlaAspIle..... 333
924 GCCGACGCGCTCTACCTCAATAGCTGCATCACCGCGCGCGGACG 973
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
334 ....AsnProSerAspAlaGlnValLeuHisLeuArgAsnIleThr 349
974 ACATGCGGGCATGTACATCTGCTTGGCGCAACACATGAGCTACAG 1023
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
349 IuAspAlaGlnGluTyrTrpCysLeuAlaGlyAsnSerIleGlyLeu 365
1024 TTCCGACGCGCTTCTCTACCGCTGCTGCGACACCAACCGCAGG 1073
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
366 HisGlnSerIleThrLeuThrValLeuProAsnGlu.....As 378
1074 ACCTGCGGCTCTGCGCTGCGGCACTACGCGCGTGGCGCGTGTCA 1123
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
378 PheLeuGlnGlnAlaGlnProAlaGlnSerArgTyrMetAspIleI 395
395 IeYTrHisSerGlyPheLeuAlaValAlaMetAlaIleMetIleVal 411
1174 CTTTGCCAGGCGCCAG 1188
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
412 LeucysArgMetGln 416
seq_name: p1R2:JC4058
seq_documentation_block:
fibroblast growth factor receptor-4 precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 29-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 20-Jun-2000
C:Accession: JC4058
R:Shiozaki, C.; Tashiro, K.; Asano-Miyoshi, M.; Saigo, K.; Emori, Y.; Shiozawa, K.
Gene 152, 215-219, 1995
A:Title: Cloning of cDNA and genomic DNA encoding fibroblast growth factor rece
A:Reference number: JC4058; MUID:95137391
A:Accession: JC4058
A:Molecule type: mRNA
A:Residues: 1-818 <SH1>
A:Cross-references: DDBJ:D31761; NID:g809527; PIDN:BA06539.1; PID:g809528
C:Genetics:
A:Introns: 43/1, 125/1, 158/1, 213/1, 254/1, 318/1, 364/1, 429/1, 478/1, 556/1, 619/1
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; pr
C:Keywords: ATP; growth factor receptor; transmembrane protein
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-818/Product: fibroblast growth factor receptor-4 #status predicted <MAT>
F:56-110/Domain: immunoglobulin homology <IM1>
F:132-137/Domain: acidic #status predicted <AD1>

```

F:177-238/Domain: immunoglobulin homology <IM2>  
 F:276-347/Domain: immunoglobulin homology <IM3>  
 F:382-402/Domain: transmembrane #status predicted <TM>  
 F:477-762/Domain: protein kinase homology <KIN>  
 F:485-493/Region: protein kinase ATP-binding motif

alignment\_scores:  
 Quality: 453.50 Length: 396  
 Ratio: 1.914 Gaps: 10  
 Percent Similarity: 59.848 Percent Identity: 29.293

alignment\_block:  
 US-09-598-042A-2 x JC4058 ..

Align seg 1/1 to: JC4058 from: 1 to: 818

```

112 GCGCCGACATATGCGCGTGCAGTCCGACGTGGAGGGGAGCCGCGCGCGT 161
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
56 GlyAsnAlaLeuArgLeuPheCysAsp..... 64
162 GACCATGTGACCAAGATGGCGCCACCATCCACAGCGGCTGAGCCGCT 211
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
65 .....ThrAsnGlnSerAsnSerIleAsnTrpTyrArgGlnGln 78
212 TCCGCGTGTGCGCGCAGGG.....CTGAG 237
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
78 sPaArgLeuLeuProGlyLysIleArgMetValGlyThrValLeuGln 94
238 GTGAGACGAGTGGAGCGGAGATGCGCGCGTGCAGCGTGCAGAGCGCAC 287
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
95 ValSerAspValThrTyrGlnAspSerGlyLeuTyrIleCys..ValVal 110
288 CAACGCGTGGCGAGCGCTGAGCGCTCACTACACCCCTGCTGCTGATG 337
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
110 LArgGlyThrClyLysIleLeuArgArgPheSerIleSerValValAsp 127
338 ACATTACCGCAGGAGAGAGCGCTGGGGCCGACAGCTCTCTGGGGGT 387
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
127 eRLeuAlaSerGlyAspGlnGlu.....AspAspGlnAspGlyArg 140
388 CAAGAGACCGCGCGCAGG.....CAGTGGC 416
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
141 ArgGlnAspThrThrAlaAspIleAsnGlnGluProValTyrPhePheG 157
417 ACAGCCCGCGCTTACACAGCGCTCCAGATGAGCGCGCGGTGATCGCAC 466
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
157 nAlaProTyrTrpThrGlnProHisArgMetAspLysLysIleHisAlaV 174
467 GCGCCGTGGTAGCTCCGTGGCGCTCAAGTGCCTGGCCAGCGGCGACCT 516
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
174 aProAlaGlyAsnThrValLysPheArgCysProAlaGlyGlySerPro 190
517 CGCGCCGACATCAGTGGATGAGAGAGACGACGCGCTTGAGC..... 558
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
191 LeuProThrIleArgTyrLeuLysAsnGlyArgGluPheArgGlyGln 207
559 .CGCCAGAGCGCGCTGAGCGCCAGAGAGAGAGAGAGAGAGAGAGAGCTGA 607
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
207 sArgIleGlyLysIleArgLeuArgHisGlnHisTrpSerLeuValMetG 224
608 AGAAGCTGGCGCGGAGAGACGCGCAATATACAGCTGCGCGTGGAGAC 657
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
224 lAsnValValProSerAspArgGlyAsnTyrThrCysValValGlnAsn 240
658 CGCGCGGCGCGCATCAAGCCACCATACAGAGTGTGATGATCCAGCGGAC 707
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
241 ArgValGlySerLeuThrTyrThrTyrPheLeuAspValLeuGlnArgSe 257
708 CCGTTCCAGCGCGTGTACAGAGACGACGACCGCTGAGACAGAGCGTGG 757
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
257 rSerHisArgProIleLeuGlnAlaGlyLeuProAlaAsnThrThrAlaA 274

```

```

758 ACTTGGGGGAGCAGCGTCTTCCATGCAAGTGCAGAGCGAGCTGAG 807
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
274 rGValGlySerAspValGlnPheTyrCysLysValTyrSerAspAlaGln 290
808 CCGGTGATCCAGTGGCTGAGAGCGGTGAGTACGGCGCGAGGGCGGCA 857
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
291 ProHisIleGlnTrpLeuLysHisIleGln..... 300
858 CAATCCACCATGATGTGGCGCGCCAGAGTT..... 891
301 .....ValAsnGlySerArgPheGlyProAspAspPhe 312
892 .....GTGCTGTGCGCCAGCGTGCAGTGTGTCGCGCGCGCGAGCG 933
312 rGlyValGlnValLeuLysThrAlaAspIle.....AsnThr 324
934 TCTACCTCAATTAAGTGTCTATCAGCCGCGCGCCAGAGAGATGCGGG 983
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
325 SerGlnValGlnValLeuHisLysLeuArgAsnIleThrMetGlnAspAlaG 341
984 CATGTACATGTGCTGTGGCGCCAGACACATGCGGTACAGCTTCCGACG 1033
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
341 yGluTyrThrCysLeuAlaGlyAsnSerIleGlyLeuSerHisGlnSerA 358
1034 CCTTCCTCAGCGTGTGCGCAGACCCAAACCGCCAGGCGCACATGTGGCC 1083
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
358 lArgPheLeuThrValLeuSerAsnGlu.....AspPheLeuGln 370
1084 TCTCTGTCTCTGCGCCACTATGACCTGCGCGCGTGCATGCGCATGCC 1133
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
371 GlnAlaGluProAlaGlnLysArgTyrMetAspIleIleTyrThrSe 387
1134 AGCGGCGCGTGTCTTATCTGAGCAGCCCTGCTGCTGCTGGCTTGGCAGG 1183
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
387 rGlyPheLeuAlaValAlaMetAlaIleValIleValValLeuCysArgM 404
1184 CCCAGAGAGACCGTGCACCGCCGCGCGCTGCCCTGCC 1221
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
404 eGlnThrProHisSerLysGlnThrLeuGlnProPro 416
seq_name: p1r2:B56182

```

seq\_documentation\_block:  
 fibroblast growth factor receptor 2 isoform IIIC - African clawed frog (fragment)  
 C/Species: Xenopus laevis (African clawed frog)  
 C/Date: 19-Oct-1995 #sequence\_revision 19-Jan-1996 #text\_change 16-Jul-1999  
 C/Accession: B56182  
 R/Mathieu, M.; Kiefer, P.; Mason, I.; Dickson, C.  
 J. Biol. Chem. 270, 6779-6787, 1995  
 A/Title: Fibroblast growth factor (FGF) 3 from Xenopus laevis (XFGF3) binds with high  
 A/Reference number: A56182; M0ID:95204476  
 A/Accession: B56182  
 A/Status: preliminary; not compared with conceptual translation  
 A/Molecule type: mRNA  
 A/Residues: 1-480 <MAT>  
 C/Superfamily: Basic fibroblast growth factor receptor 1; immunoglobulin homology; pr  
 C/Keywords: growth factor receptor  
 F:183-244/Domain: immunoglobulin homology <IM>

alignment\_scores:  
 Quality: 453.00 Length: 409  
 Ratio: 1.880 Gaps: 13  
 Percent Similarity: 58.924 Percent Identity: 30.073

alignment\_block:  
 US-09-598-042A-2 x B56182 ..

Align seg 1/1 to: B56182 from: 1 to: 480

```

112 GCGCGACATATGCGCGTGCAGTGCAGTGGAGGGGAGCCGCGCGCGT 161
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
71 GlyGluProLeuAspLeuArgCysProLeuAlaAspGly.....ProLe 85

```

```

1023 CTTCCGACAGCGCTTCCTCACCGAGTCGGCCAGGACCAAAACCGCCAGGGC 1072
      | :::::::::::::::::::::: | | | | |
362 rGlnHisSerIatIrrLeuThrValHisProAlaProValAsnPro.... 377
      :::::::::::::::::::::: ::::::::::::::
1073 CACCTGTGGCCTCTCTCTCCGCGCCACTAGCCGCGCGTGGCCGGTGC 1122
      :: :: :: :: :: :: :: :: :: :: ::
378 .....LeuGlnAspAsnProValIrrProTyrTyrIrrMetGln 388
      ::::::::::::::::::::::
1123 ATGCGCATGCCAGCGCGCGCTGTTCATCCCTGGGACCGCTGCTGTG 1172
      | | | | | | | | :: :: :: :: :: :: :: :: :: :: ::
389 IlegIyIetIyrSerThrGlyIlePheHeIleIrrPheCysMetValVal... 404
      :::::::::::::: ::::::::::::::
1173 GCTTGGC.....CAGGCCCAAGAAAGCCCT 1198
      | | | | | :: :: :: :: :: :: :: :: :: :: ::
405 'ValGlyValAlcysArgMetArgGlnGlyAlaLysIleLysLysAsn 421
      :::::::::::::: ::::::::::::::
1199 GCACCCCGCGCGCTGCCCGCTCCCGG.....CGTGGGCAACCGC 1236
      | | | | | | | | :: :: :: :: :: :: :: :: :: :: ::
421 heThrGlyProProValHisLysLeuThrLysArgIleProLeuHisArg 437
      :::::::::::::: ::::::::::::::
1237 CGCGCGGGAGCGCGCGCGCGAGCGAGC 1263
      :::::::::: :: :: :: :: ::
438 GlnValThrValSerAlaAspSerSer 446

```

receptor precursor - human

N/Contains: protein-tyrosine kinase (EC 2.7.1.112)

C;Species: Homo sapiens (man)

C;Date: 16-Sep-1992 #sequence\_revision 16-Sep-1992 #text\_change 16-Jul-1999  
C;Accession: 315335

C/ACCESSION: 310230  
R:Seno, M.: Sasada, R.: Watanabe, T.: Ishimaru, K.: Tsurashi, K.

Biochim. Biophys. Acta 1089, 244-246, 1991

A; Title: Two cDNAs encoding novel human FGF receptor  
1; Reference number: 010005 with 0100456

A:Accession: S16236  
A:Reference number: S16236; MUID:912/4336

A;Molecule type: mRNA

A;Residues: 1-769 &lt;SEN&gt;

References: EMBL: X56191; NID: g29431; PID: g29431; A; Cross-references: EMBL: X56191; NID: g29431; PID: g29431; C: Superfamily: basic fibroblast growth factor receptor 1; Immunoglobulin

**C; Keywords:** ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor; immunoglobulin homology; protein kinase C; signal transduction; tyrosine phosphorylation.

```
F;1-21/Domain: signal sequence #status predicted <SIG>
```

F:22-378/Domain: extracellular #status predicted <EVT>  
F:22-69/Product: fibroblast growth factor receptor #3

F;132-138/Region: acidic

F;172-233/Domain: immunoglobulin homology &lt;IMM&gt;

Accession	Protein	TMpred	TMHMM	PhyS	PhyS2	PhyS3	PhyS4	PhyS5	PhyS6	PhyS7	PhyS8	PhyS9	PhyS10	PhyS11	PhyS12	PhyS13	PhyS14	PhyS15	PhyS16	PhyS17	PhyS18	PhyS19	PhyS20	PhyS21	PhyS22	PhyS23	PhyS24	PhyS25	PhyS26	PhyS27	PhyS28	PhyS29	PhyS30	PhyS31	PhyS32	PhyS33	PhyS34	PhyS35	PhyS36	PhyS37	PhyS38	PhyS39	PhyS40	PhyS41	PhyS42	PhyS43	PhyS44	PhyS45	PhyS46	PhyS47	PhyS48	PhyS49	PhyS50	PhyS51	PhyS52	PhyS53	PhyS54	PhyS55	PhyS56	PhyS57	PhyS58	PhyS59	PhyS60	PhyS61	PhyS62	PhyS63	PhyS64	PhyS65	PhyS66	PhyS67	PhyS68	PhyS69	PhyS70	PhyS71	PhyS72	PhyS73	PhyS74	PhyS75	PhyS76	PhyS77	PhyS78	PhyS79	PhyS80	PhyS81	PhyS82	PhyS83	PhyS84	PhyS85	PhyS86	PhyS87	PhyS88	PhyS89	PhyS90	PhyS91	PhyS92	PhyS93	PhyS94	PhyS95	PhyS96	PhyS97	PhyS98	PhyS99	PhyS100	PhyS101	PhyS102	PhyS103	PhyS104	PhyS105	PhyS106	PhyS107	PhyS108	PhyS109	PhyS110	PhyS111	PhyS112	PhyS113	PhyS114	PhyS115	PhyS116	PhyS117	PhyS118	PhyS119	PhyS120	PhyS121	PhyS122	PhyS123	PhyS124	PhyS125	PhyS126	PhyS127	PhyS128	PhyS129	PhyS130	PhyS131	PhyS132	PhyS133	PhyS134	PhyS135	PhyS136	PhyS137	PhyS138	PhyS139	PhyS140	PhyS141	PhyS142	PhyS143	PhyS144	PhyS145	PhyS146	PhyS147	PhyS148	PhyS149	PhyS150	PhyS151	PhyS152	PhyS153	PhyS154	PhyS155	PhyS156	PhyS157	PhyS158	PhyS159	PhyS160	PhyS161	PhyS162	PhyS163	PhyS164	PhyS165	PhyS166	PhyS167	PhyS168	PhyS169	PhyS170	PhyS171	PhyS172	PhyS173	PhyS174	PhyS175	PhyS176	PhyS177	PhyS178	PhyS179	PhyS180	PhyS181	PhyS182	PhyS183	PhyS184	PhyS185	PhyS186	PhyS187	PhyS188	PhyS189	PhyS190	PhyS191	PhyS192	PhyS193	PhyS194	PhyS195	PhyS196	PhyS197	PhyS198	PhyS199	PhyS200	PhyS201	PhyS202	PhyS203	PhyS204	PhyS205	PhyS206	PhyS207	PhyS208	PhyS209	PhyS210	PhyS211	PhyS212	PhyS213	PhyS214	PhyS215	PhyS216	PhyS217	PhyS218	PhyS219	PhyS220	PhyS221	PhyS222	PhyS223	PhyS224	PhyS225	PhyS226	PhyS227	PhyS228	PhyS229	PhyS230	PhyS231	PhyS232	PhyS233	PhyS234	PhyS235	PhyS236	PhyS237	PhyS238	PhyS239	PhyS240	PhyS241	PhyS242	PhyS243	PhyS244	PhyS245	PhyS246	PhyS247	PhyS248	PhyS249	PhyS250	PhyS251	PhyS252	PhyS253	PhyS254	PhyS255	PhyS256	PhyS257	PhyS258	PhyS259	PhyS260	PhyS261	PhyS262	PhyS263	PhyS264	PhyS265	PhyS266	PhyS267	PhyS268	PhyS269	PhyS270	PhyS271	PhyS272	PhyS273	PhyS274	PhyS275	PhyS276	PhyS277	PhyS278	PhyS279	PhyS280	PhyS281	PhyS282	PhyS283	PhyS284	PhyS285	PhyS286	PhyS287	PhyS288	PhyS289	PhyS290	PhyS291	PhyS292	PhyS293	PhyS294	PhyS295	PhyS296	PhyS297	PhyS298	PhyS299	PhyS300	PhyS301	PhyS302	PhyS303	PhyS304	PhyS305	PhyS306	PhyS307	PhyS308	PhyS309	PhyS310	PhyS311	PhyS312	PhyS313	PhyS314	PhyS315	PhyS316	PhyS317	PhyS318	PhyS319	PhyS320	PhyS321	PhyS322	PhyS323	PhyS324	PhyS325	PhyS326	PhyS327	PhyS328	PhyS329	PhyS330	PhyS331	PhyS332	PhyS333	PhyS334	PhyS335	PhyS336	PhyS337	PhyS338	PhyS339	PhyS340	PhyS341	PhyS342	PhyS343	PhyS344	PhyS345	PhyS346	PhyS347
-----------	---------	--------	-------	------	-------	-------	-------	-------	-------	-------	-------	-------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------

E:480-765/Domain: protein kinase homology <KIN>  
F:480-765/Domain: intracellular #status predicted

F;488-496/Region: protein kinase ATP-binding motif

F;62-107,179-231,278-340/Disulfide bonds: #status predicted

F:518,535,627/Active site: Iys, Glu, Asp #status predicted  
F:83,123,228,241,263,291,318,329,350/Binding site: carbonhy

[illegible]

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

alignment\_scores: 452.50

Ratio:	1.926	Gaps:	10
--------	-------	-------	----

Percent Similarity: 61.039 Percent Identity: 30.909

alignment block:

US-09-598-042A-2 x S16236 .

1000

Align seg 1/1 to: \$16236 from: 1 to: 769

100 GTGGCCCGGCTGGGCCCGCACTATGCGGGCTGCAGTGCCCACTGGAGGGGGA 149

[illegible]

51 ValAlaAlaProGlyGluSerLeuGluValArgCysLeuLeuLys..... 65

150 CCCGCCCGCTGACCATGTGGACCAAGGATGCCCGACCATGCCACAGCG 199

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

```

66 .AspAlaAlaValIleSerTrpThrLysAspGly.....ValHisLeug 80
200 GCTGGACCGCTTCCGGCTGCTG...CCGACGGGCTGAGGTAGACAG 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
80 LyrProAsnAsnArgThrValLeuIleGlyLysTrpLeuGlnIleLysGly 96
247 GTGAGCGCGGAGATGCCGGGCTAGCTGTGCAAGGCCACCAACGGCTT 296
    : : : : : : : : : : : : : : : : : : : : : : : : : : :
97 AlathrProAlaGAspSerGlyLeuThrAlaCysThrAlaSerArgThrVa 113
297 CGCGACGCTGAGCGTCACTACACCTGTCGTGCTGATGACATTAAGCC 346
    ||| : : : : : : : : : : : : : : : : : : : : : : :
113 LAspSerGluThrTrpTrpThrMetValAsnValThrAspAlaIleSer 130
347 CAGGAGAGAGAGCGCTGGGCGCCGACAGCTCTCTGGGGGTCAAGAGAC 396
    ||| : : : : : : : : : : : : : : : : : : : : : : :
130 erGlyAspAsp.....GluAspAspThrAspGlyAlaGluAsp 142
397 CCGCGCCAGCCAG.....CAGTGGGACAGCGCGCTTCACACAGCC 437
    : : : : : : : : : : : : : : : : : : : : : : : : :
143 PheValSerGluAsnSerAsnAsnLysArgAlaProTyrTrpThrAsnTh 159
438 CTCCAAGATGAGCGCGCGGTGATCGACGCGCGGTAGTCCGCTGC 487
    : : : : : : : : : : : : : : : : : : : : : : : : :
159 rGluLysMetGlyLysArgLeuHisAlaValProAlaAlaAsnThrValL 176
488 GGCCTAAGTGGTGGCCAGCGGGGACCCCTGGCCGACATCATCGTGAATG 537
    : : : : : : : : : : : : : : : : : : : : : : : : :
176 ySPheArgCysProAlaGlyLysAsnProMetProThrMetArgTrpLeu 192
538 AAGGACAGACAGCGCTGAGC.....CGCCACAGAGCGCGCTGAGCC 578
    ||| : : : : : : : : : : : : : : : : : : : : : : :
193 LysAsnLysLysGluLysLysGlnHisArgLleGlyLysLysVal 209
579 CAGGAGAGAGAGTGAACACTGACCTGAGAACCTCGCGCGGAGAGACA 628
    ||| : : : : : : : : : : : : : : : : : : : : : : :
209 LArgAsnGlnHisTrpSerLeuIleMetGluSerValValProSerAspL 226
629 GCGCAAAATACACCTGCGCGGTGTCGAAACCGCGCGCGCCATCAGCC 678
    : : : : : : : : : : : : : : : : : : : : : : : : :
226 ySGLysAsnTyrThrCysValValGluAsnGlnTyrGlySerLleAsnHis 242
679 ACCATCAAGGAGTGTGATCCACGAGCCCGTCCAAAGCCCGGCTGAC 728
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
243 ThrTyrHisLeuAspValValGluArgSerProHisArgProIleLeuG1 259
729 AGGACACCGCCCGCTGAAACAGACGCTGACTTGGGGGAGCACAGTCTT 778
    : : : : : : : : : : : : : : : : : : : : : : : : :
259 nAlaGlyLeuProAlaAsnAlaSerThrValValGlyLysPValGluP 276
779 TCCAGTCCAAAGTCCGACGACGATGAAAGCCGATGATCCAGTGGCTGAAG 828
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
276 heValCysLysValTyrSerAspAlaGlnProHisIleGlnTrpLleLys 292
829 CGGCTGAG.....TACGGCGCCGAGGGCGCCGCCCAACTC 863
    : : : : : : : : : : : : : : : : : : : : : : : : :
293 HisValIleLysAsnGlySerLysTyrGlyProAspGly..... 305
864 CACCATGATGTGGCGCGCAGAAATTGTGTGCTGCCCGCCAGCGGTACG 913
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
306 .....LeuProTyrLeuLysV 311
914 TG.....TGTCGGGCGCGACGCTCCATCAATTAAGCTGCTGATC 957
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
311 alLeuLysHisSerGlyIleAsnSerSerAsnAlaGluValLeuAlaLeu 327
958 ACCGTCGCGCGCCGACGAGATGGGCGCATGTACTGCTGCTGGCGGCA 1007
    : : : : : : : : : : : : : : : : : : : : : : : : :
328 PheAsnValThrGluAlaAspAlaGlyLysLysLysValSerAs 344
1008 CACCATGGGTACAGCTTCCGACGCGCTTCTACAGTGTGCGCAGAC 1057
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
344 nTyrIleGlyGlnAlaAsnGlnSerAlaIleTrpLeuThrValLeuProLysG 361

```

```

1058 CAAAACCGCCAGGCGCACCTGTGGCTTCCTGCTCGGCGACATAGCCTG 1107
    : : : : : : : : : : : : : : : : : : : : : : : : :
361 LngAlaIleProGly.....ArgGluLysGluIleThrAlaSer 373
1108 CCGTGGCGCGCGTGCATCGGATCCAGCGCGCGCTGCTTCATCTGAGG 1157
    ||| : : : : : : : : : : : : : : : : : : : : : : :
374 ProAspTyrLeuGluIleAlaIleTyrCysIleGlyValPheLeuIleAl 390
1158 CACCTGCTCTGCTGCTTTCGACGCGCCAGAGAGAGCGCTGACCCCGG 1207
    : : : : : : : : : : : : : : : : : : : : : : : : :
390 acyMetValValThrValIleLeuCysArgMetLysAsnThrThrLysL 407
1208 CGCCT 1212
    |||
407 ySPPro 408
seq_name: p1r2:A45081
seq_documentation_block:
  fibroblast growth factor receptor 2 IIb, FGFR2 IIb - human
  N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
  C:Species: Homo sapiens (man)
  C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999
  C:Accession: A45081
  R:Deil, K.R.; Williams, L.T.
  J. Biol. Chem. 267, 21225-21229, 1992
  A:Title: A novel form of fibroblast growth factor receptor 2. Alternative splicing of
  A:Reference number: A45081; M0ID:93016048
  A:Accession: A45081
  A:Status: preliminary
  A:Molecule type: nucleic acid
  A:Residues: 1-822 <DEL>
  A:Cross-references: GB:M97193; NID:g182566; PIDN:AAA52449.1; PID:g182567
  A:Experimental source: placenta
  A:Note: sequence extracted from NCBI backbone (NCBIP:116207)
  C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; pr
  C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor recept
  F:172-233/Domain: Immunoglobulin homology <IMM>
  F:379-399/Domain: transmembrane #status predicted <TM>
  F:480-765/Domain: protein kinase homology <KIN>
  F:488-496/Region: protein kinase ATP-binding motif
  F:62-107,179-231,278-340/Disulfide bonds: #status predicted
  F:518,535,627/Active site: Lys, Glu, Asp #status predicted
alignment_scores:
  Quality: 452.50      Length: 385
  Ratio: 1.926        Gaps: 10
  Percent Similarity: 61.039      Percent Identity: 30.909
alignment_block:
  US-09-598-042A-2 x A45081 ..
Align seg 1/1 to: A45081 from: 1 to: 822
100 GTGGCCCGCGGTGGCGCGCATGCGGCTGCGAGTCCCGATGAGGGGGA 149
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
51 ValAlaIleAlaProGlyLysSerLeuGlnValArgCysLeuLeuLys..... 65
150 CCGCGCGCGCTGACATGTGGACCAAGATGGCGCGCACATCCACAGCG 199
    : : : : : : : : : : : : : : : : : : : : : : : : :
66 .AspAlaIleValIleSerTrpThrLysAspGly.....ValHisLeug 80
200 GCTGAGCGCGCTTCCGGCTGCTG...CCGACGGGCTGAGGTAGACAG 246
    ||| : : : : : : : : : : : : : : : : : : : : : : :
80 LyrProAsnAsnArgThrValLeuIleGlyLysTrpLeuGlnIleLysGly 96
247 GTGAGCGCGGAGATGCCGGGCTAGCTGTGCAAGGCCACCAACGGCTT 296
    : : : : : : : : : : : : : : : : : : : : : : : : :
97 AlathrProAlaGAspSerGlyLeuThrAlaCysThrAlaSerArgThrVa 113
297 CGCGACGCTGAGCGTCACTACACCTGTCGTGCTGATGACATTAAGCC 346

```

```

113 1aspserglutntrptprphemevalasnvalthrspalallesters 130
347 CAGGGAGAGAGCCCTGGGGCCGACAGCTCCTGGGGGTCAAGGAGC 396
130 erglyaspasp.....gluaspsprhsplalaglyasp 142
397 CCGCCAGCCAGCAG.....CAGTGGGCACAGCCGCTTCACACAGCC 437
143 phevastsergluasnerasnlnslyargalaprotyltriphrtsnrh 159
438 CTCGAAGATGAGCGCCGGGTGATCCACGCGCCGTGGTAGTCCGTGC 487
159 rdluymsatglulysargleuhsalavalproalalaaasnthvall 176
488 GGGTCAGGCGTGGCCAGCGGACACCTGGCGCCGACATCAACGTGATG 537
176 ysbheargcysproalaglygluasnprometprothrmetcargtrpleu 192
538 AAGCAGCAGCAGCCCTTGACG.....CGCCAGAGCCGCTGAGCC 578
193 lysasnlglyllysglnphelysglnghlnhsarglleglytyrlysva 209
579 CAGGAAGAAGAAGTGACACACTGACCTGAAGAACTGCGCGCGGAGACA 628
209 laryasnghlnstprserleuilemetgluservalvalproseraspl 226
629 GGGGCAATATACACTGGCCGCTGTGACACCGCGCGCGCCATCAACGCC 678
226 ysglyasnlythrlycysvalvalgluasnglulutyglyserlleasnhs 242
679 ACCTCAAGATGATGTGATCCAGCGCCGCTTCACACCCGCTGCTCAC 728
243 thrlyrhlnhsleuaspvalvalgluargserprhnhsargprolleucl 259
729 AGCAGCAGCAGCCCGTGAACACAGCGGTGACTTCGGGGAGACACGCTT 778
259 nalagllyleuproalalasnalasethrvalvalglylyasprvalglup 276
779 TCCAGTGCAGAGTGGCCGACGAGCGTGAACCGTGAATCCAGGCGTAG 828
276 hevalyllysvaalytyrseraspalaglnprohlnstlleglntrplelys 292
829 CCGCTGGAC.....TAGCGCGCGCAGGCGCGCCACAACTC 863
293 hlsvalgllyasnglyserlytyrlylproaspgly..... 305
864 CACCATCATGTGGCGCGCCAGAACTTGTGCTGCTCCCGACGGGTGACG 913
306 .....leuprotyrleuylsv 311
914 TG.....TGGTGGGGCGCCGAGCGCTCTACCTCAATAGAGCTGCATC 957
311 albeulyshtsersgllytleasnserasnalaaglvalleualaleu 327
958 ACCCGTCCCGCAGCAGCAGTGCATGATCATCTGCTTGGCGCA 1007
328 pheasnvalthrglualaspralaglylutyrllecylsvalseras 344
1008 CACCATGCGCTTACAGCTTCGCGACGCGCTTCTCACGTGTCGACAGCC 1057
344 ntyrlleglyglinalaasnghlnserlatrpleuthrvalleuprolys 361
1058 CAACACGCGGCGCAGCTGTGGCTCTCGTGGTCCGCGCCTACACCTG 1107
361 lnglnlaprogly.....arglulysglululethralaser 373
1108 CCGTGGCCCGCTGATCGCATCCAGCGCGCGCTTTCATCTCGTGG 1157
374 proasprtyrleuaglulalealalytyrlylleglyvalpheuleucl 390
1158 CACCGTCTCTGTGGCTTGGCGCCAGAGAAAGCCGTGCACCCCG 1207

```

```

390 acysmetvalvalthrvalilleuencysargmetlyasntrhrlysl 407
1208 CGGCT 1212
407 ysPro 408
seq_name: p1r2:A41794

```

```

seq_documentation_block:
keratinocyte growth factor receptor precursor - human
N:Alternate names: fibroblast growth factor receptor-related protein K-SAM
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C:Accession: A41794; A44775
R:Miki, T.; Bottaro, D.P.; Fleming, T.P.; Smith, C.L.; Burgess, W.H.; Chan, A.M.; Aar
Proc. Natl. Acad. Sci. U.S.A. 89, 246-250, 1992
A:Title: Determination of ligand-binding specificity by alternative splicing: two dis
A:Reference number: A41794; MUID:92108030
A:Accession: A41794
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-822 <MTK>
A:Cross-references: GB:M80634; NID:g186740; PIDN:AAA36147.1; PID:g186741
A:Note: sequence extracted from NCBI backbone (NCBI:P:75096)
R:Champion-Arnaud, P.; Ronsin, C.; Gilbert, E.; Gesnel, M.C.; Housaint, E.; Breathna
Oncogene 6, 979-987, 1991
A:Title: Multiple mRNAs code for proteins related to the B6K fibroblast growth factor
A:Reference number: A44775; MUID:91296403
A:Accession: A44775
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 263-362 <CH4>
A:Cross-references: GB:S04853
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; pr
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor recept
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-822/Product: keratinocyte growth factor receptor #status predicted <MAT>
F:172-233/Domain: immunoglobulin homology <IMM>
F:379-399/Domain: transmembrane #status predicted <TM>
F:480-765/Domain: protein kinase homology <KIN>
F:488-496/Region: protein kinase ATP-binding motif
F:62-107,179-231,278-340/Dissulfide bonds: #status predicted
F:518,535,627/Active site: Lys, Glu, Asp #status predicted

```

```

alignment_scores:
Quality: 452.50 Length: 385
Ratio: 1.926 Gaps: 10
Percent Similarity: 61.039 Percent Identity: 30.909

```

alignment\_block:

US-09-598-042a-2 x A41794 ..

Align seg 1/1 to: A41794 from: 1 to: 822

```

100 GTGGCGCGCTGGCGCGACATAGCGGTGATGCCAGTGGAGGGGA 149
||||| |||:|||||:|||||:|||||:|||||:|||||:|||||:
51 ValAlaAlaProGlyGlySerleuGluValArgCysLeuLeuLys..... 65
150 CCGCGCGCGCTGACCATGTGACCAAGAGTGGCGCGACCATCCACAGCG 199
66 .aspAlaAlaValIleSerIleThrLysaspLy.....ValHisLeuG 80
200 GCTGAGCGCGCTTCGCGGTGCTG...CCGACGGCGCTGAAGTGAAGCAG 246
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
80 LyrAsnAsnArGThrValLeuIleGlyGlyTyrLeuGlnIleLysGly 96
247 GTGGAGCGGAGGATCCGCGGTGATGCTGTGCAAGCGCACCAACGCGCT 296
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
97 AlaThrProArgAspSerGlyLeuTyrAlaCysThrAlaSerArgGlyVa 113
297 CGGACCGCTGAGCGCTCACTACACCCCTGCTGTGCTGATGCAATTAGCC 346

```



```

113 1ASPSeGIuThrTrpTyrlPheMetValAsnValThrsPAlaIleSers 130
134 CAGGAGAGAGACCTGGGGCCGACAGCTCTGGGGGTCAAGAGAC 396
130 eRGIyAspAsp.....GIuAspAspThrsPcLYAlaGIuAsp 142
397 CCCGCCAGCCAG.....CAGTGGGCACGACCGCGCTTCAACAGCC 437
143 PheValSerGIuAsnSerAsnAsnLysArgAlaProTYrTrpThrsnth 159
438 CTCAGAGTGGGGCCGGGATGACAGCGCCGTGGTGGTCCCGGC 487
159 rGIuLysMetGIuLysArgLeuHISAlaValProAlaIleAsnthrVal 176
488 GGCCTCAAGTGGCGTGGCCAGCGGACCTCGCCGACATACAGTG 537
176 ySPheATGcYsProlaGlyLysnProMetProthrMetArgTrpleu 192
538 AAGACGACACAGCGCTTGACG.....CGCCAGAGCGCGCTGAGCC 578
193 LysAsnGIyLysGIuPheLysGIuHISArgIleGIyLysVal 209
579 CAGGAGAGAGAGTGGAGACACAGCTGACAGACCTGGCGCGGAGCA 628
209 LArgAsnGIHIStrPserLeuIleMetGIuSerValAlaProSerAspL 226
629 GCGGCAGATACACCTGGCGCTGTGAGACCGCGGCGCGCCATCAACGCC 678
226 ySGIyAsnTrTyrlThcYsValAlaGIuAsnGIuTYrGlySerIleAsnHIS 242
679 ACCTACAGAGTGGATGTGATCCAGCGGACCGCTTCCAGACCGCTGAC 728
243 ThTYrHISLeuAspValAlaGIuArgSerProHISArgProIleLeuGI 259
729 AGGACGACGACCGCGTGAACAGAGTGGGACTGGGGGGAGAACCGTCT 778
259 nAlaGIyLeuProAlaAsnAlaSerThValValGIyLysAspAlaGIuP 276
779 TCCAGTGCAGAGTGGCGACGAGCTGAGAGCGGTGATCCAGTGGCTGAG 828
276 nValGIyLysValLysSerAspAlaGIuProHISIleGIuTrpleLys 292
829 CGCGTGCAG.....TACGCGCCGAGGCGCGCCACACATC 863
293 HISValGIuLysAsnGIySerLysTrGIyProAspGIy..... 305
864 CACCATGATGTGGCGGCGGACAGATTGTGTGCTGCCACAGGCTGACG 913
306 .....LeuProTYrLeuLysV 311
914 TG.....TGTCGCGCGCCGACGCTCTACCTCAATACCTGCTCATC 957
311 AlLeuLysHISerGIyLISAsnSerAsnAlaGIuValLeuAlaLeu 327
958 ACCGTCGCGCCGACGAGCATGGGCGATGATCATCTCTGGCGGCA 1007
328 PheAsnValThrGIuAlaAspAlaGIyLysTrIleCysLysValSerAs 344
1008 CACCATGGGCTACAGCTCCGACAGCGCTTCTCAGCTGCTGCCAGAC 1057
344 nTYrTleGIyAlaIleAsnGIuSerAlaTrpleuThrValLeuProLys 361
1058 CAAGACGCGAGGCGACCTGTGGCTCTCTGCTCGGCACATGAGCTG 1107
361 InGIuAlaProGIy.....ArgGIuLysGIuIleThrAlaSer 373
1108 CCGTGGCGCGGTGCATCGGATCCAGCGCGCGCTGTCTTCACTGGG 1157
374 ProAspTYrLeuGIuIleAlaIleTYrCysIleGIyValPheLeuIleAl 390
1158 CACCTGTCTCTGTGGCTTGGCAGGCGGACAGAGCCGCTGACCCCG 1207

```

```

390 acYsMetValValThrValIleLeuCysArgMetLysAsnthrTrpLysL 407
1208 CGCCT 1212
407 ySPro 408
seq_name: p1r2:S17295
seq_documentation_block:
  fibroblast growth factor receptor - mouse
  C:Species: Mus musculus (house mouse)
  C:date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 16-Jul-1999
  C:Accession: S17295
  A:title: PCR-based identification of new receptors: molecular cloning of a receptor f
  A:Reference number: S17295; MWID:91270892
  A:Accession: S17295
  A>Status: preliminary
  A:Molecule type: mRNA
  A:Residues: 1-820 <RAZ>
  C:Cross-references: EMBL:X55441; NID:950141; PIDN:CAA39083.1; PID:950142
  C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; pr
  F:478-763/Domain: protein kinase homology <KIN>

```

```

alignment_scores:
  Quality: 447.50      Length: 378
  Ratio: 1.946        Gaps: 9
  Percent Similarity: 60.847      Percent Identity: 31.217

```

```

alignment_block:
  US-09-598-042A-2 x S17295 ..

```

```

Align seg 1/1 to: S17295 from: 1 to: 820

```

```

118 ACTATGCGGCTGCAGTGGCCAGTGGAGGGGACCGCGCGCTGACCAT 167
57 SerLeuGIuLeuGIuCysMetLeuLys.....AspAlaIleAlaIleSe 71
168 GTGAGCAGAGATGGCGCGCACATCCAGCGGCTGGAGCGCGCTCCGG 217
71 TrTrpTrpLysAspGIy.....ValHISLeuGIuProAsnAsnArgThrv 86
218 TGCTG...CCGACGCGGCTGAAGGTGAAGCAGAGTGGAGCGGAGATGCC 264
86 AlLeuIleGIyArgTYrLeuGIuHISLeuLysGIyAlaIleThrProArgSpsr 102
265 GCGGTGTACGTGTGCAAGGCCACACAGCGCTTGGCAGCGCTGAGCGTCA 314
103 GIyLeuTYrAlaCysThrAlaIleArgThrValAlaSerGIuThrTrpTy 119
315 CTACACCTCGTCGTGTGATGATGATTAAGCCAGGAGGAGAGACCTGG 364
119 rPheMetValAsnValThrsPAlaIleSerSerGIyAspArgIuAspRA 136
365 GCGCCAGAGCTCTGTGGGGGTCAAGAGAGCCCGGACGAGAGTGG 414
136 sPThAspSerSerGIuArgValSerGIuAsnArgSerAsnIleArg... 151
415 GCACGACGCGGCTTGCACACAGCGCTTCAAGATGAGCGCGGGGTATGCG 464
152 ..AlaProTYrTrpThrsnthrGIuLysMetGIuLysArgLeuHISAl 167
465 ACGGCGCGGTGTACCTCGTCGCGGCTCAAGTGCCTGGCCAGCGGAC 514
167 aValProAlaIleAsnthrValLysPheArgCysProAlaGIyLysAsp 184
515 CTGGCGCGGATCAGCTGATGAAGAGACGACGACCGCTTGACG..... 558
184 rOThrProThrMetArgTrpleuLysAsnGIyLysGIuPheLysGIuInGI 200

```



```

776 CCTTCAGTGCAGTGGCGGACGAGCGGATGATCCAGTGGCTG 825
    :||| ||||| ||||| ||||| ||||| ||||| |||||
277 IupreValCysValTyrSerAspAlaGlnProHisIleGlnTrpIle 293
    :||| ||||| ||||| ||||| ||||| ||||| |||||
826 AAGCGCGTGGAG.....TACGGCGCCGAGGCGCCGACAA 860
    :||| ||||| ||||| ||||| ||||| ||||| |||||
294 LysHisValGluArgAsnGlySerTyrClyProAspGly..... 307
    :||| ||||| ||||| ||||| ||||| ||||| |||||
861 CTCACCATCATGTGGCGCCGAGAAATTGTGTGCTGCCACGGGTG 910
    :||| ||||| ||||| ||||| ||||| ||||| |||||
308 .....LeuProTyrLeuG 312
    :||| ||||| ||||| ||||| ||||| ||||| |||||
911 ACGTG.....TGGTCGGCGCCGACGGCTCTACCTCAATAAGCTGCTC 954
    :||| ||||| ||||| ||||| ||||| ||||| |||||
312 InValLeuLysHisSerGlyTyrLeuSerSerAsnAlaGluValLeuThr 328
    :||| ||||| ||||| ||||| ||||| ||||| |||||
955 ATCACCCTGGCGCCGACGAGATGCGGCGATGACATCTGCTTGGCGC 1004
    :||| ||||| ||||| ||||| ||||| ||||| |||||
329 LeuTyrAsnValThrGluAlaAspAlaGlyLysTyrIleGlyLysValSe 345
    :||| ||||| ||||| ||||| ||||| ||||| |||||
1005 CACACACATGGGCTACAGCTTCCGACGCGCTTCCCTCACGCTGCCAG 1054
    :||| ||||| ||||| ||||| ||||| ||||| |||||
345 rAsnTyrIleGlyGluAlaAsnGlnSerAlaTrpLeuSerValLeuProS 362
    :||| ||||| ||||| ||||| ||||| ||||| |||||
1055 ACCCAAAACCGCGGCGACCTGTGGCTCTGCTCTGGCGCACATAGC 1104
    :||| ||||| ||||| ||||| ||||| ||||| |||||
362 erLeuGlnAlaPro.....GluLysGluLysGluPheProThr 374
    :||| ||||| ||||| ||||| ||||| ||||| |||||
1105 CTGCGCTGGCGCCGTGATCGGATCGGATCCAGCGCGCTGTCTTACATCT 1154
    :||| ||||| ||||| ||||| ||||| ||||| |||||
375 SerProAspTyrLeuGlnIleAlaIleTyrCysIleGlyValPheLeuI 391
    :||| ||||| ||||| ||||| ||||| ||||| |||||
1155 GGGACACCTGCTCTGCTGGCTTCCGACGCGCCAGAAAGCCGTGCACC 1204
    :||| ||||| ||||| ||||| ||||| ||||| |||||
391 eAlaCysMetValLeuThrValIleLeuCysArgMetLysAsnThrThr 408
    :||| ||||| ||||| ||||| ||||| ||||| |||||
1205 CCGGGCGCT.....GCCCTGCC 1221
    :||| ||||| ||||| ||||| ||||| ||||| |||||
408 yslYsProAspPheSerSerGlnProAlaValHisLysLeuThrLysArg 424
    :||| ||||| ||||| ||||| ||||| ||||| |||||
1222 CTGCGTGGCGACCGCGCGCGCGGACCGCGCGACGACGACGAGAC.. 1269
    :||| ||||| ||||| ||||| ||||| ||||| |||||
425 IleProLeuArgArgGlnValThrValSerAlaAspSerSerSerMe 441
    :||| ||||| ||||| ||||| ||||| ||||| |||||
1270 .....A 1270
    :||| ||||| ||||| ||||| ||||| ||||| |||||
441 tAsnSerAsnThrProLeuValArgIleThrThrArgLeuSerSerThrA 458
    :||| ||||| ||||| ||||| ||||| ||||| |||||
1271 AGACACTTCCCTGCTGGCGCCCTCAGCGCTGGCGCTGTGGGGCTG 1320
    :||| ||||| ||||| ||||| ||||| ||||| |||||
458 lAspAlaProMetLeuAlaGlyValSer..... 467
    :||| ||||| ||||| ||||| ||||| ||||| |||||
1321 TGTGAGAGCATGGCTCTCGCGACGCCGCCAG 1353
    :||| ||||| ||||| ||||| ||||| ||||| |||||
468 .....GluTyrGluLeuProGluAspProLys 476
    :||| ||||| ||||| ||||| ||||| ||||| |||||

seq_name: p1r2:sl8209

seq documentation block:
  fibroblast growth factor receptor 4 precursor (clone 61) - mouse
  N:Alternate names: tyrosine kinase Mpk-11
  N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
  C:Species: Mus musculus (house mouse)
  C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 16-Jul-1999
  C:Accession: S18209; S26751; S30497
  R:Stark, K.L.; McMahon, J.A.; McMahon, A.P.
  Development 113, 641-651, 1991
  A:Title: FGFR-4, a new member of the fibroblast growth factor receptor family, expressed
  A:Reference number: S18209; MUID:92146274
  A:Accession: S18209
  A:Molecule type: mRNA
  A:Residues: 1-799 <STA1>
  A:Cross-references: EMBL:X59927

```

```

R:Stark, K.L.
submitted to the EMBL Data Library, May 1991
A:Reference number: S26751
A:Accession: S26751
A:Molecule type: mRNA
A:Residues: 1-485, 'OVRAEARG', 486-799 <STA2>
A:Cross-references: EMBL:X59927; NID:950968; PIDN:CAA4251.1; PID:950969
R:Gillardt-Hebenstreit, P.; Nieto, M.A.; Frain, M.; Mattei, M.G.; Chestier, A.; Wilkin
Oncogene 7, 2499-2506, 1992
A:Title: An Eph-related receptor protein tyrosine kinase gene segmentally expressed 1
A:Reference number: S30496; MUID:93096484
A:Accession: S30497
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 611-667 <GIL>
A:Cross-references: EMBL:X57236; NID:953187; PIDN:CAA40512.1; PID:953188
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; pr
C:Keywords: ATP; duplication; glycoprotein; growth factor receptor; phosphotransferas
F:1-18/Domain: signal sequence #status predicted <Sig>
F:19-799/Product: fibroblast growth factor receptor 4 #status predicted <MAT>
F:19-366/Domain: extracellular #status predicted <EXT>
F:162-223/Domain: immunoglobulin homology <IMM>
F:367-387/Domain: transmembrane #status predicted <TM>
F:388-799/Domain: intracellular #status predicted <INT>
F:462-747/Domain: protein kinase homology <KIN>
F:470-478/Region: protein kinase ATP-binding motif
F:54-98, 169-221, 268-330/Disulfide bonds: #status predicted
F:500, 517, 609/Active site: Lys, Glu, Asp #status predicted

alignment_scores:
  Quality: 442.00      Length: 455
  Ratio: 1.720        Gaps: 23
  Percent Similarity: 56.484      Percent Identity: 32.527

alignment_block:
US-09-598-042A-2 x S18209  ..

Align seg 1/1 to: S18209 from: 1 to: 799

109 CTGGGCGCCAGTATGCGGCTGACAGTCCAGTGAGGGAGACCGCGCC 158
    ||||| :||| ||||| ||||| ||||| ||||| |||||
46 LeuGlyGlnProValArgLeuCysCys.....GlyArgThrGluArg 59
    ||| ||||| ||||| ||||| ||||| ||||| |||||
159 GCTGACCATGTGACACAGATGGCGGCACCATCCACAGC..... 198
    ||| ||||| ||||| ||||| ||||| ||||| |||||
59 gGlyArgHisTrpTyrLysGluGlySerArgLeuAlaSerAlaGlyArgv 76
    ||| ||||| ||||| ||||| ||||| ||||| |||||
199 .....GGCTGAGCCGCTTCCGCGTGGCTGCGACGGGCTCAAGTGAAG 243
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
76 aArgGlyTrp...ArgGlyArg.....LeuGluIleAla 86
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
244 CAGGTGAGCGGGAGATGCGCGGTGACGTAGTGCAGGACCAACAGG 293
    :||| ||||| ||||| ||||| ||||| ||||| |||||
87 SerPheLeuProGluAspAlaGlyArgTyrLeuCysLeuAlaArg..... 101
    :||| ||||| ||||| ||||| ||||| ||||| |||||
294 CTGCGCAGCCTGAGCGTCACTACACCTGCTGCTGCTGATGAC.... 339
    ||||| :||| ||||| ||||| ||||| ||||| |||||
102 ....GlySerMetThrValValHisAsnLeuThrLeuLeuMetLysAspS 117
    :||| ||||| ||||| ||||| ||||| ||||| |||||
117 erLeuThrSerIleSerAsnAspGluAspProLysThrLeuSerSerSer 133
    :||| ||||| ||||| ||||| ||||| ||||| |||||
379 TCTGGGGGTCAAGAGAGACCGCGCCAGCGATGGCGACGACGCGCTT 426
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
134 SerSerGlyHisValTyrPro.....GlnGln.....AlaProTyrTr 146
    :||| ||||| ||||| ||||| ||||| ||||| |||||
429 CACACAGCCTTCAGATGAGAGCGCGGATGATGACAGCGCGCTGAGTA 478
    :||| ||||| ||||| ||||| ||||| ||||| |||||
146 pThrHisProGlnArgMetGluLysLysLeuHisAlaValProAlaGlyA 163
    :||| ||||| ||||| ||||| ||||| ||||| |||||
479 GCTCGTGGCGCTCAAGTGCCTGGCGACGGGACACCTTGGCGCGACATC 528
    :||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

163 ..:::|||||:.....:::|||||
164 snthralValysPheargCysProAlaCysArgAsnProMetProthIle 179
529 ACCTGATGAGGACGACGACGCTTG.....ACGGCCCAAGAGC 569
180 H1strPleuLysAspLysAlaPheN1sgLysAsnArgIleGly 196
570 CGCTGAGCCGAGAGAGAGAGAGTGGACCTGAGACCTGAAGAACCTGCGG 619
196 ValLeuArgLeuArgHisGlnHisTrpSerLeuValMetGluSerValP 213
620 CGGAGGACGACGCGCAAAATACACTGCGCGGTGTCGACCGCGCGGCGC 669
213 roSerAspArgGlyThrTyThrCysLeuValGlnAsnSerLeuGlySer 229
670 ATCAAGCCGACCTCAAGAGTGGATGTGATCCAGCGGACCGCTCCAGCC 719
230 IleArgTySerTyLeuLeuAspValLeuGluArgSerProHisArgP 246
720 CGTGTCAAGGACGACCGCGCTGAACAGACGAGTGGACTTGCGGGGGA 769
246 oIleLeuGlnAlaGlyLeuProAlaAsnThrThrAlaValAlGlySer 263
770 CCAAGTCTTCAGTCAAGTGGACGCGACGAGCTGAAGCGGTGATCCAG 819
263 sPValGluLeuLeuCysLysValTySerAspAlaGlnProHisIleGln 279
820 TGGCTGAAGCGCTGAGTACGCGGCGGCGGCGGCGCACACATCCACAT 869
280 TrPLeuLysHisValIle.....AsnLysSerSerP 291
870 CGATGTGGGCGCCAGAGTTTGTG...GTGTCGCCAGGAGTACGTG 916
291 eGlyAlaAspGlyPheProTyValGlnValLeuLysThrThrAspIle 307
917 GGTGGGCGCGGACGCTCTACCTCAATAGCTGTCATACCGCTGCC 966
308 .....AsnIleSerGluValGlnValLeuTyLeuAsnVal 320
967 CGCGAGGACGATGCGGCGATGTACATCTGCTGGCGGCGCAACATCGG 1016
321 SerAlaGlnAspAlaGlyGlyTyThrCysLeuAlaGlyAsnSerIleG 337
1017 CTACAGCTTCGCGAGCGCTTCCTCACCGCTGCCAGACCCCAAAACCG 1066
337 yLeuSerTyGlnSerAlaTrpLeuThrValLeuProGlu..... 350
1067 CAGGGCACCTGTGGCTCTGCTCGGCTAGCCCTAGCCCTGG... 1113
351 .....GluAspLeuThrTrpThr 356
1114 .....CCGGTGCATCGGCAATCCGAC 1136
357 ThrAlaThrProGluAlaArgTyThrAspIleLeuTyValSerG 373
1137 CGGCGCTGTCTTCATCGTGGGACCGCTGCTCTGTGC..... 1174
373 ySerLeuValLeuLeuValLeuLeuLeu..AlaGlyValTyHisArg 389
1175 ...TTTGCCAGGCGCCAGAGAACCGCTGCACCC..... 1204
390 GlnValIleLeuArgLysTySerArgGlnProValThrIleGlnLysI 406
1205 .CGCGGCTGCGCCCTCGCTGCGGCGACCGCGCGGCGGAGCGCGG 1253
406 uSerArgPheProLeu.....AlaArgGln..... 414
1254 CGACGCGACGAGGACAGGACCTTCCTGTTGGCGCGCTCAGCGCTG 1303
415 ..PheSerLeuGlnSerArgSerGlyLys.....SerSerLeu 427
1304 GCCCTGTGTGGGCTGTGTGAGAGAGCATGAGTCTCGGACGCGCCGAG 1353
:::||||| ||| ||||| |||||

```

```

428 SerLeuValArgGly...ValArgLeuSerSer.SerGly...ProProL 442
1354 CACTTACTGGG 1364
442 euLeuThrGly 445
seq_name: p1r2:A56182

```

```

seq documentation block:
C:Species: Xenopus laevis (African clawed frog)
C:Date: 19-Oct-1995 #sequence_revision 19-Jan-1996 #text_change 16-Jul-1999
C:Accession: A56182
R:Mathieu, M.; Kiefer, P.; Mason, I.; Dickson, C.
J. Biol. Chem. 270, 6779-6787, 1995
A:Title: Fibroblast growth factor (FGF) 3 from Xenopus laevis (XRF3) binds with high
A:Reference number: A56182; MUID:95204476
A:Accession: A56182
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-480 <MAT>
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; pr
C:Keywords: growth factor receptor
F:183-244/Domain: Immunoglobulin homology <IMM>

```

```

alignment_scores:
Quality: 438.00 Length: 413
Ratio: 1.848 Gaps: 14
Percent Similarity: 57.385 Percent Identity: 29.782

```

```

alignment_block:
US-09-598-042a-2 x A56182 ..

```

```

Align seg 1/1 to: A56182 from: 1 to: 480

```

```

112 GCGCGCATATGGCGCTGAGTCCAGTGGAGGGGACCGCGCGCT 161
||||| : : |||||
71 GlyLupProLeuAspLeuArgCysProLeuAlaAspGly.....ProLe 85
162 GACCATGTGACCAAGATGGCGCGCACATCCACAGCGGCTGAGCGCT 211
||||| |||||
85 uValThrTrpThrLysAspLysAlaLysLeuGluValAsn...AsnArg 101
212 TCGCGCTGTCGCGGCGGAGTGAAGTGAAGCGTGAAGCGAGAGAT 261
||||| : : |||||
101 hLeuLeuValArgThrTyLeuGlnIleLysGluSerThrProAlaAsp 117
262 GCGCGGTGATGTCGACAGGCGCACCAAGCGCTTCGACGCTGAGCGT 311
||||| : : |||||
118 SerGlyLeuTyAlaCysSerValLeuLysAsnSerHisPhePheHisVa 134
312 CACATACACCTGCTGCTGATGATACATTAAGCCAGGAGAGAGCGC 361
||||| |||
134 LAsnValThr..... 137
362 TGGGCGCCACAGCTCTCTGGGGGTCAAGAGAGC.....CCGCGCAGC 405
||||| |||||
138 ....GluAlaSerSerSerGlyAspAspLysAsnAspAsnArgLys 152
406 CAGCAGTGGGACAGA.....CCGCGCTTCACACA 434
||||| : : |||||
153 GluAspPheThrAsnAspAsnAsnIleArgAlaProTyTrpHisAs 169
435 GCGCTCAAGATGAGCGCGGTGATTCGACGCGCGCGTGGAGTCCG 484
: : |||||
169 nThrGluLysMetGlnLysLysLeuHisAlaValProAlaAlaAsnThrV 186
485 TGGCGCTCAAGTGTGTCGACGCGGCGACCTTCGCGCGCACATCAGTGG 534
||||| : : |||||
186 alLysLeuArgCysProAlaGlyLysAsnProThrProArgMetArgTrp 202
535 ATGAAGACGACGACGCGCTTGACG.....CGCCAGAGGCGCGCTGA 575

```



```

697 ATCCAGCGAGCCGCTTCAAGCCGCTGCTACAGCAGCCACCCCGTGA 746
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
248 VAGLVARSerProHisProIleLeuGlnAlaGlyLeuProAlaAs 264
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
747 CACGACGTGGAGCTTGGGGGACACAGCTCTCCAGTCAGACGTGGCA 796
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
264 nfhtrhtrhlysvaGlyGlyNspArgGluPheValCysLysValTyrS 281
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
797 GCGACGTGAAGCCGGTGAATCCAGTGGCTGAAGCGCTGGAGTACGGCGC 846
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
281 etrspAlaGlnProHisIleGlnTTP..... 289
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
847 GAGGCGCGCCACAATCCACATCATGATGGGGCGCCAGAAAGTTGTGG 896
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
290 .....ArgHis.....IleGluLeuAsnGlySerLysIle..... 299
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
897 GCTGCCACGAGGTGACGTGTGTGGCGCCGACGCGCTCTACCTCAATA 946
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
300 .....GlyProAspGlyAsnProTyrLeu 308
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
947 AGCTGCTC..... 954
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
308 ysvallLeuysAlaAlaGlyAsnThrThrValLysGluIleGluVal 324
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
955 .....ATCACCCGTCGCCGCCAGAGATGCGGCATGATCATCTGCT 998
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
325 LeuTyrValArgAsnValSerPheGluAspAlaGlyGluTyrThrCys 341
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
999 TGGCGCCCAACACCATGGCTTACAGCTTCGACGCGCTTCTCATCCGTGC 1048
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
341 uAlaGlyAsnSerIleGlyIleSerTyrHisSerAlaTrpLeuThrVal 358
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
1049 TGGCAGACCCCAAAACCGCCAGCGCACCTGTGCTCTCTGCTGCGGCC 1098
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
358 euProAspGluGluArgGlnLeuAsp.....SerSerSerSerGluTyr 372
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
1099 ACTAGCGCTGCCGTGGCCCGCTGTATCGGCATCCAGCGCGCTGTCTT 1148
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
373 ThrGluIle.....AlaIleTyr 378
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
1149 CATCTGGCGACCCCTGCTCTG 1170
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
378 rCysValGlyGlyPheLeuIle 385
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
seq_name: p1r1.TVCHFG
seq_documentation_block:
Fibroblast growth factor receptor 1 precursor - chicken
N:Alternate names: basic fibroblast growth factor receptor
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) cekl
C:Species: Gallus gallus (chicken)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C:Accession: A11345; A33908
R:Lee, P.L.; Johnson, D.E.; Cousins, L.S.; Fried, V.A.; Williams, L.T.
Science 245, 57-60, 1989
A:Title: Purification and complementary DNA cloning of a receptor for basic fibroblast g
A:Reference number: A11345; MUID:89298406
A:Accession: A11345
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-819 <LE>
A>Note: part of the sequence was confirmed by protein sequencing
R:Pasquale, E.B.; Singer, S.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 5449-5453, 1989
A:Title: Identification of a developmentally regulated protein-tyrosine kinase by using
A:Reference number: A33908; MUID:89315814
A:Accession: A33908
A:Molecule type: mRNA
A:Residues: 1-89, 'A', '91-685, 'W', 687-819 <PAS>
A:Cross-references: GB:M24637
A>Note: this protein is expressed in embryonic tissues and, at low levels, in adult brai
C:Genetics:
A:Gene: cekl

```

```

C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; pr
C:Keywords: Arp; autophosphorylation; duplication; glycoprotein; growth factor recept
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-819/Product: fibroblast growth factor receptor 1 #status predicted <MAT>
F:22-374/Domain: extracellular #status predicted <EXT>
F:125-132/Region: acidic
F:169-230/Domain: immunoglobulin homology <IMM>
F:375-395/Domain: transmembrane #status predicted <TM>
F:396-819/Domain: intracellular #status predicted <INT>
F:474-759/Domain: protein kinase homology <KIN>
F:482-490/Region: protein kinase ATP-binding motif
F:54-100,176-228,275-339/Disulfide bonds: #status predicted
F:76,116,225,238,262,294,315,328/Binding site: carbohydrate (Asn) (covalent) #status
F:512,529,621/Active site: Lys, Glu, Asp #status predicted
F:626,639/Binding site: magnesium (Asn, Asp) #status predicted
F:652/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi
alignment_scores:
      Quality: 434.00      Length: 443
      Ratio: 1.625      Gaps: 13
Percent Similarity: 60.271      Percent Identity: 27.765
alignment_block:
US-09-598-042a-2 x TVCHFG ..
Align seg 1/1 to: TVCHFG from: 1 to: 819
103 GCCCGCTGGCGCGCTATGCGGCTGCAAGTCCCAAGTGGAGGGAGCC 152
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
44 AlaHisProGlyAspLeuLeuGlnLeuArgCysArgLeuAspAspAsp 60
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
153 GCGCGCGCTGACCATGTGAGCCAGAGGATGCGCCACATCAACAGCGGT 202
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
60 LginSerIleAsn..TrpValArgAspGlyValGlnLeuProGluAsn 75
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
203 GGAAGCGCTTCCGCGCTGCTGCCAGGCGCTGAAGTGAAGAGGTGGAG 252
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
76 ..AsnArgThrArgIleThrGlyGluGluValGluValArgAspArgVal 91
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
253 CGGAGAGATGCCGCGCTGTACGTGTCAAGGCCACACAGCGCTTGGCAG 302
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
92 ProGluAspSerGlyLeuTyrAlaCysMetThrAsnSerProSerGlySe 108
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
303 CCTGACGTCACACTACACCCCTGCTGCTGGATGACATTAGCCAGGA 352
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
108 r...GluThrThrTyrPheSerValAsnValSerAspAlaLeuProSerA 124
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
353 AGGAGAGCTGGGCGCCAGACGCTCTGGGGGTCAAGAGACCCGCC 402
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
124 IacLusAspAspAspArgGluAspAspSerSerSerGluGluValAla 140
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
403 AGC.....CAGCAGTGGCAGACGCGCTTACACAGCCCTC 440
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
141 AspAsnThrLysProAsnGlnAlaValAlaProTyrTrpThrTyrProG 157
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
441 CAAGATGAGGCGCGGCTGATCGCAGCGCCGCGGTAGCTCCGCGC 490
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
157 uLysMetGluLysLysLeuHisAlaValProAlaAlaLysThrValLysP 174
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
491 TCAATGCTGTGGCCAGCGGACCCCTCGGCCGACATCACTGATGAAG 540
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
174 helysCysProSerGlyGlyThrProAsnProThrLeuArgTrpLeuLys 190
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
541 GACGACAGAGCTTGAAGCGCCAGAG.....GCGCTAGCC 578
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
191 AsnGlyLysGluPhe...LysProAspHisArgIleGlyGlyTyrTyrS 206
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
579 CAGGAAGAAGAGTGGACACTGACCTGAAGAACGCGCGCGAGAGACA 628
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
206 IArgTyrAlaThrTrpSerIleIleMetAspSerValValProSerAsp 223
      :::::|||||:| | | | | | | | | | | | | | | | | | | |
629 GCGCAATATACACTTCCGCTGTGACACCGCGCGCCCATCAACGCC 678
      :::::|||||:| | | | | | | | | | | | | | | | | | | |

```

[illegible]

```
seq_documentation_block:
  fibroblast growth factor receptor 2 precursor - African clawed frog
  N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
  C:Species: Xenopus laevis (African clawed frog)
  C>Date: 21-Jan-1994 #sequence_revision 27-Oct-1995 #text_change 16-Jul-1999
  C:Accession: A49123; S25060
  R:Friesel, R.; Brown, S. A.
  Development 116, 1051-1058, 1992
```

```

A:Title: Spatially restricted expression of fibroblast growth factor receptor-2 during
A:Reference number: A49123; MUID:93201992
A:Accession: A49123
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-813 <FR1>
A:Cross-references: EMBL:X65943
A:Note: sequence extracted from NCBI backbone (NCBI:P.128003)
R:Brown, S.N.; Friesel, R.E.
submitted to the EMBL Data Library, May 1992
A:Description: Xenopus fibroblast growth factor receptor-2: cloning, expression and d
A:Reference number: S25060
A:Accession: S25060
A:Molecule type: mRNA
A:Residues: 1-632, 'A', 634-813 <BRO>
A:Cross-references: EMBL:X65943; NID:664694; PIDN:CAA6758.1; PID:664695
C:Function:
A:Description: receptor mediating fibroblast growth factor roles in patterning of dev
A:Note: expressed from embryonic stage 11 through tadpole stages
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; pr
C:Keywords: ATP; autophosphorylation; duplication; embryo; glycoprotein; growth facto

F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-813/Product: fibroblast growth factor receptor 2 #status predicted <MAT>
F:18-367/Domain: extracellular #status predicted <EXT>
F:124-130/Region: acidic
F:163-223/Domain: immunoglobulin homology <IMM>
F:368-388/Domain: transmembrane #status predicted <TMN>
F:389-813/Domain: intracellular #status predicted <INT>
F:469-754/Domain: protein kinase homology <KIN>
F:477-485/Region: protein kinase ATP-binding motif
F:58-103,170-221,268-332/Disulfide bonds: #status predicted
F:79,115,231,255,287,308,321/Binding site: carbohydrate (Asn) #status pred
F:507,552,616/Active site: Lys, Glu, Asp #status predicted
F:651,653/Binding site: magnesium (Asn, Asp) #status predicted
F:647/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi

alignment_scores:
Quality: 425.50 Length: 409
Ratio: 1.780 Gaps: 14
Percent Similarity: 58.435 Percent Identity: 29.584

alignment_block:
US-09-598-042A-2 x A49123 ..

Align seg 1/1 to: A49123 from: 1 to: 813

```

```

112 GCGCGCCTATGCGGGCTGCAGATGCCAGTGGAGAGGGGAGACCGCGCGCGCT 161
      ||||:  ::  |||||:||||:||||:
51 G|G|L|P|ro|leu|asp|leu|arg|G|Ser|Pro|leu|a|a|sp|gly|.....Pro|le 65
162 GACCATGTGGACCAAGATAGGGCGACCATTCACAGCGGGCGTGAAGCGGCT 211
      ||||: |||||:|||||||  :::::  ::  ::|||
65 u|a|l|t|h|T|P|T|P|T|L|y|a|s|p|o|l|y|a|l|a|l|y|s|e|u|l|u|a|s|n|..|A|s|n|A|T|G|T 81
212 TCGCGCGCTGCGCGGAGGGCGTGAAGGTGAACAGTGTGAGCGGAGGAT 261
      ::||:  ::||:  ||||:||||:||||:||||:
81 h|r|e|u|l|e|v|a|l|g|T|h|T|T|T|T|y|l|e|u|G|l|n|l|e|y|G|l|u|S|e|T|T|T|T|A|T|A|G|A|S|P 97
262 GCGCGCGCTGTACGTGTGCAAGGCCAACCAACGGCTTCGGCAGCCTGAGCGT 311
      ::||:||||:||||:||||:||||:  :::::  ::  ||
98 S|e|r|C|l|y|e|u|T|l|a|c|y|S|e|r|V|a|l|l|e|u|L|y|a|s|n|S|e|r|T|h|S|P|h|e|N|h|S|A 114
312 CAACATCAACCCCTGCTGCTGCTGGATGACATTATGCCCAAGGAAAGAGACC 361
      ||||  |||
114 l|a|s|n|v|a|l|T|h|.....  117
362 TGGGGCCCCGACATCCTCTGGGGGGTCAAGAGAGAC.....CCGGCCAGC 409
      |||||:|||||  ::|||
118 .....G|u|a|l|a|S|e|r|S|e|r|G|l|y|a|s|p|a|S|p|l|u|a|s|p|a|S|n|a|S|p|o|l|y|S|e|r 133
406 C|A|C|A|G|T|G|G|G|C|A|C|A|G|A|.....CCGCGCTTCACACA 434

```





[illegible]

heparin-binding growth factor receptor variant alpha-a2 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 28-Feb-1992 #sequence revision 28-Feb-1992 #text-change 16-Jul-1999  
 C:Accession: C40862  
 R:Hon, J.; Kan, M.; McKeenan, K.; McBride, G.; Adams, P.; McKeenan, W.L.  
 Science 251, 665-668, 1991  
 A>Title: Fibroblast growth factor receptors from liver vary in three structural domains  
 A:Reference number: A40862; MUID:91126480  
 A:Accession: C40862  
 A:Status: Preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-662 <HOU>  
 A:Cross-references: GB:M63888  
 C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protein  
 C:Keywords: ATP: growth factor receptor; heparin binding; transmembrane protein  
 F:171-232/Domain: immunoglobulin homology <IIM>  
 F:377-399/Domain: transmembrane #status predicted <TM>  
 F:476-618/Domain: protein kinase homology #status atypical <KIN>  
 F:484-492/Region: protein kinase ATP-binding motif

alignment_scores:	Quality:	423.50	Length:	442
Ratio:	1.580	Gaps:	14	
Percent Similarity:	60.633	Percent Identity:	27.602	
alignment block:				
US-09-598-042A-2 x C40862 ..				
Align seg 1/1 to: C40862 from: 1 to: 662				
112	GGCCGCACTAGTGGCGCTGAGGCGCCATGTGGAGGGGAGACCCGCCGCCGT	161	GGCGGCGTGTACGTGTGCACAGGCCACACAGCGCTTCCGACCTGACGCT	311
48	GLysApLeuLeuGlnLeuAlaArgCysArgLeuAlaArgAspAspValGlnSerI	64	GLysApLeuLeuGlnLeuAlaArgCysArgLeuAlaArgAspAspValGlnSerI	64
162	GACCAATGTGTGACCAAGGATGTGGCCGACCATCCACAGCGCTGAGCCGCT	211	GACCAATGTGTGACCAAGGATGTGGCCGACCATCCACAGCGCTGAGCCGCT	211
64	eAsn..TrpLeuAlaArgAspGlyValGlnLeuAlaGlnSer..AsnArgT	79	eAsn..TrpLeuAlaArgAspGlyValGlnLeuAlaGlnSer..AsnArgT	79
212	TCGCGTGTCTGCCGAGGCGCTGAAGGTGAACAGAGTGCAGCGGAGAGAT	261	TCGCGTGTCTGCCGAGGCGCTGAAGGTGAACAGAGTGCAGCGGAGAGAT	261
79	hArgLLeuThrGlyGlnGluValGlnValGlnAspSerValProAlaAsp	95	hArgLLeuThrGlyGlnGluValGlnValGlnAspSerValProAlaAsp	95
262	GCGCGCGTGTACGTGTGCACAGGCCACACAGCGCTTCCGACCTGACGCT	311	GCGCGCGTGTACGTGTGCACAGGCCACACAGCGCTTCCGACCTGACGCT	311
96	SerGlyLeuTyrValAlaCysValThrSerSerProSerGlySer..AspTh	111	SerGlyLeuTyrValAlaCysValThrSerSerProSerGlySer..AspTh	111
312	CACATTAACCCCTGTGCTGTGATGACATTAGCCACAGGAAGAGAGACC	361	CACATTAACCCCTGTGCTGTGATGACATTAGCCACAGGAAGAGAGACC	361
111	rThrTyrPheSerValAsnValSerAspAlaLeuProSerSerGlnAspA	128	rThrTyrPheSerValAsnValSerAspAlaLeuProSerSerGlnAspA	128
362	TGGGGCCGACAGCTCTGTGGGGTCAAGAGACCCGCC.....	402	TGGGGCCGACAGCTCTGTGGGGTCAAGAGACCCGCC.....	402
128	sPAspAspAspAspAspAspSerSerSerGlnGlyLysGlnThrAspAsnThr	144	sPAspAspAspAspAspAspSerSerSerGlnGlyLysGlnThrAspAsnThr	144
403	.....AGCCAGCACTGGGACAGCCGCGCTTACACAGCCCTCCAGAT	446	.....AGCCAGCACTGGGACAGCCGCGCTTACACAGCCCTCCAGAT	446
145	LysProAsnAlaArgMetProValAlaProTyrLTrpThrSerProGlnLysMe	161	LysProAsnAlaArgMetProValAlaProTyrLTrpThrSerProGlnLysMe	161
447	GAGCGCCGCGGTGATGCGACGCGCCGCTGGTAACTCGTGGCGTCAAGT	496	GAGCGCCGCGGTGATGCGACGCGCCGCTGGTAACTCGTGGCGTCAAGT	496
161	fcuLysLysLeuAlaValAlaValProAlaAlaLysThrValLysPheLysC	178	fcuLysLysLeuAlaValAlaValProAlaAlaLysThrValLysPheLysC	178
497	GGGTGGCGAGCGGAGACCTCGCGCCGACATCAGTCAGTAAAGAGAGAC	546	GGGTGGCGAGCGGAGACCTCGCGCCGACATCAGTCAGTAAAGAGAGAC	546
178	ysProSerSerGlyThrProAsnProThrLeuAlaTyrTrpLeuLysAsnGly	194	ysProSerSerGlyThrProAsnProThrLeuAlaTyrTrpLeuLysAsnGly	194
547	CAGCGCTTACCGCCGCGCAG.....GCCCGTGAAGCCACAGAA	584	CAGCGCTTACCGCCGCGCAG.....GCCCGTGAAGCCACAGAA	584
195	LysGlnPhe..LysProAspHisArgLLeuGlyLysValAlaGly	210	LysGlnPhe..LysProAspHisArgLLeuGlyLysValAlaGly	210
585	GAAGAAGTGACACTGAGCCTGAGACCTGCGCGGAGAGACAGCGCA	634	GAAGAAGTGACACTGAGCCTGAGACCTGCGCGGAGAGACAGCGCA	634

```

210 rAlaThrTpSerIleIleMetAspSerValValProSerAspLysGlyA 227
635 AATACACCTCCGCGTGCATACCGGCGGCATACACGCACCTAC 684
:::|||||:::|||||:::|||||:::|||||
227 smYrThrCysIleValGluAsnGluTyrGlySerIleAsnHisThrTyr 243
685 AAGGTGATGTGATCCAGCGACCCGCTCCAGCCGCTGCTACAGGAC 734
:::|||||:::|||||:::|||||:::|||||
244 GluLeuAspValValGluArgSerProHisArgProIleuGluIleAsp 260
735 GCACCCCGTGAACACGAGCTGACCTCCGGGGGACACCTCTCCACT 784
|||||:::|||||:::|||||:::|||||
260 yLeuProAlaAsnLysThrValAlaLeuGlySerAspValGluPheMetC 277
785 GCAAGTGGCCAGCAGCTGATCCAGCCGCTGATCCAGTGGTGAAGCGGTG 834
|||||:::|||||:::|||||:::|||||
277 yslYsValTyrSerAspProGluProHisIleGluThrPheLysHisIle 293
835 GAGTACGGCGCCGAGGGCCGACCAACTCCACCATGATGGGCGGCCA 884
|||:::|||||:::|||||:::|||||
294 GluVal.....AsnGlySerLysIleGlyProAspAsnLe 305
885 GAAGTTTGTG..GTGTCGCCACAGGGTGACGTGTGTCGGCGCCGACG 931
:::|||||:::|||||:::|||||
305 uProTyrValGluIleLeuLysThrAlaGlyVal.....AsnThrThra 330
932 GCTCTACATCAATTAAGCTGCTATCATCACCCTCCGCGCCAGCATGCG 981
:::|||||:::|||||:::|||||
320 spLysGluMetGluValLeuHisLeuArgAsnValSerPheGluAspAla 336
982 GGCATGATACATGCTGCTGGCGCCACACCATGAGGTACAGCTCCGACG 1031
|||:::|||||:::|||||:::|||||
337 GlyGluTyrThrCysLeuAlaGlyAsnSerIleGlyLeuSerHisHisSe 353
1032 CGGCTTCCTACCGCTGTGTGCCA.....GACCCAAACCCCGCAGGGCCAC 1075
|||||:::|||||:::|||||
353 rAlaThrPheThrValLeuIleAlaLeuGluGluArgPro..... 366
1076 CTGTGGCTCTGCTGCTCGCCGACATGAGCTGCGGCGCCGCTGTCATC 1125
:::|||||:::|||||
367 .....AlaValMetThrSerProLeuTyrIleGluIleIle 378
1126 GGCATCCACCGCGCGCTGCTTCATC.....CTGGGACCTGCT 1166
:::|||||:::|||||
379 IleTyrCysThrGlyAlaPheLeuIleSerCysMetValGlySerValI 395
1167 CGGTGGCTTGGCAGGCGCCAGAAAGCGG..... 1197
:::|||||:::|||||
395 eValTyrLysMetLysSerGlyThrLysLysSerAspPheHisSerGlnM 412
1198 ..TGACACCCCGCGCGCTCCCTCCCTGCGTGGGACCGCGCGCGGG 1245
:::|||||:::|||||
412 eValAlaValHisLysLeuAlaLysSerIleProLeuArgAlaValThr 428
1246 ACGCGCCGCGACCGGACGAGACAGAC.....CT 1277
:::|||||:::|||||
429 ValSerAlaLysPheSerAlaSerMetAsnSerGlyValLeuLeuValAr 445
1278 TCCTCGTGGCGCGCGCTCAGCGCTGCGCTGGTGGGGGCTGTGTGAG 1327
|||||:::|||||:::|||||
445 gProSerArgLeuSerSerSerGlyThrProMetLeuAlaGlyValSerG 462
1328 AGCATGGGTCTCCGCGACGCCCGAC 1353
|||:::|||||
462 LuTyrGluLeuProGluAspProArg 470

```

seq\_name: p1r1.TVHUF6

seq\_documentation\_block:

fibroblast growth factor receptor 1 precursor - human  
 N:Alternate names: basic fibroblast growth factor receptor; heparin-binding growth factor  
 N:Contains: HBGR 3; HBGR alpha b1; HBGR beta a1; HBGR beta b1; HBGR gamma a1; HBGR  
 C:Species: Homo sapiens (man)

C>Date: 31-Mar-1989 #sequence\_revision 31-Dec-1993 #text-change 16-Jul-1999  
 C:Accession: S11692; B40862; S09226; A61533; A61536; S26739; S26738; S17374; S17375;  
 R:Dione, C.A.; Crumley, G.; Bellor, F.; Kaplow, J.M.; Searfoss, G.; Ruta, M.; Burges  
 EMO J. 9, 2685-2692, 1990  
 A>Title: Cloning and expression of two distinct high-affinity receptors cross-reactin  
 A:Reference number: S11691; MUID:90360977  
 A:Accession: S11692  
 A:Molecule type: mRNA  
 A:Residues: 1-822 <DIO>  
 A:Cross-references: EMBL:X52833; NID:g31377; PIDN:CAA37015.1; PID:g31378  
 R:Hou, J.; Kan, M.; McKeenan, K.; McBride, G.; Adams, P.; McKeenan, W.L.  
 Science 251, 665-668, 1991  
 A>Title: Fibroblast growth factor receptors from liver vary in three structural doma  
 A:Reference number: A40862; MUID:91126480  
 A:Accession: B40862  
 A:Molecule type: mRNA  
 A:Residues: 1-822 <HOU>  
 A:Cross-references: GB:M63887  
 A>Note: potentially twelve variants may arise by alternative splicing in liver cells  
 R:Isacchi, A.; Bergonzoni, L.; Sarmientos, P.  
 Nucleic Acids Res. 18, 1906, 1990  
 A>Title: Complete sequence of a human receptor for acidic and basic fibroblast growth  
 A:Reference number: S09226; MUID:90245600  
 A:Accession: S09226  
 A:Molecule type: mRNA  
 A:Residues: 1-816, 'R', 818-822 <ISA>  
 A:Cross-references: EMBL:X51803; NID:g31367; PIDN:CAA36101.1; PID:g31368  
 R:Wenstrom, S.; Sandstrom, C.; Claesson-Welsh, L.  
 Growth Factors 4, 197-208, 1991  
 A>Title: cDNA cloning and expression of a human FGF receptor which binds acidic and b  
 A:Reference number: A61533; MUID:92118394  
 A:Accession: A61533  
 A:Molecule type: mRNA  
 A:Residues: 1-147, 150-193, 'S', 195-822 <MEN>  
 A:Cross-references: GB:M46641; NID:g182529; PIDN:AAA35835.1; PID:g182530  
 A:Experimental source: teracarcinoma cell line Tera-2  
 R:Kiefer, M.C.; Baird, A.; Nguyen, T.; George-Nascimento, C.G.; Mason, O.B.; Boley, L  
 Growth Factors 5, 115-127, 1991  
 A>Title: Molecular cloning of a human basic fibroblast growth factor receptor cDNA an  
 A:Reference number: A61536; MUID:92118399  
 A:Accession: A61536  
 A>Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-147, 150-822 <KIE>  
 R:Tronick, S.R.  
 Submitted to the EMBL Data Library, January 1991  
 A:Reference number: S19167  
 A:Accession: S26739  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-822 <TRI>  
 A:Cross-references: EMBL:X57121; NID:g31392; PIDN:CAA40403.1; PID:g31393  
 A:Accession: S26738  
 A:Molecule type: mRNA  
 A:Residues: 1-147, 150-822 <TR2>  
 A:Cross-references: EMBL:X57120; NID:g31390; PIDN:CAA40402.1; PID:g31391  
 R:Eisemann, A.; Ahn, J.A.; Graziani, G.; Tronick, S.R.; Roa, D.  
 Oncogene 6, 1195-1202, 1991  
 A>Title: Alternative splicing generates at least five different isoforms of the human  
 A:Reference number: S17373; MUID:91319400  
 A:Accession: S17374  
 A:Molecule type: mRNA  
 A:Residues: 1-30, 120-147, 150-822 <E14>  
 A:Cross-references: EMBL:X57119; NID:g31386; PIDN:CAA40404.1; PID:g31387  
 A>Note: the sequence from Fig. 3 is inconsistent with that from Fig. 1 in lacking 25-  
 A>Note: this form is designated isoform I  
 A:Accession: S17375  
 A:Molecule type: mRNA  
 A:Residues: 1-30, 120-822 <E13>  
 A:Cross-references: EMBL:X57119; NID:g31388; PIDN:CAA40401.1; PID:g31389  
 A>Note: the sequence from Fig. 3 is inconsistent with that from Fig. 1 in lacking 25-  
 A>Note: this form is designated isoform II  
 A:Accession: S17377





[illegible]

A:Cross-references: GB:M6503; NID:g19329; PIND:AAA37620.1; PID:g193293  
A:Experimental source: breast cancer cell line SC-3  
A>Note: The authors translated the codon ATG for residue 397 as Ile and C TG for resid  
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; pr  
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor recept  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-83/Product: fibroblast growth factor receptor #status predicted <FGF>  
F:181-242/Domain: immunoglobulin homology <IMM>  
F:387-409/Domain: transmembrane #status predicted <TMN>  
F:487-771/Domain: protein kinase homology <KIN>  
F:494-502/Region: protein kinase ATP-binding motif  
E:67-113,188-240,287-351/Disulfide bonds: #status predicted  
E:524,541,633/Active site: Lys, Glu, Asp #status predicted

alignment\_scores:  
Quality: 419.50 Length: 440  
Ratio: 1.565 Gaps: 15  
Percent Similarity: 60.909 Percent Identity: 28.182

alignment\_block:  
US-09-598-042A-2 x JH0393 ..  
  
Align seg 1/1 to: JH0393 from: 1 to: 832

```
112 GGCCGCACATATCGGCTGCAGTGGCCCATGTGGAGGGGAGCCCCGCCGT 161
      ||| :|||:|||||:||||| :||| :||| :|||
60 GLYAsPLeuIeUgInLeuArGcysArgLeuArGAsPValGIserIlI 76
162 GACCAGTGTGACCACGAGTAGGC...CGCACATCCACAGCGGTGTGAGCC 208
      ||| :|||:|||||:||||| :||| :||| :|||
76 eAsn...TPleuArGAsPGLyAlGIleuValGIser.....AsnA 90
209 CCTTCGCGCTGTGCGCAGGGGCTGAAGTGAAAGCAGGTGAGCGGAG 258
      || ||||:|:| :|||:|||||:||||| :||| :||| :|||
90 rGrhArGlIeThrGIygluGIuValGIValArGspSerIlIeProAla 106
259 GATGCGGGGCTGACTGTGTCAAGGCCACCAACGGTTGGCAGCCTGAG 308
      |||:|||||:|||||:||||| :|||:|||||:||||| :||| :||| :|||
107 AsPSerGIyleuTyralAcysValThrSerSerProSeGIyser...As 122
309 CGGCACATCACACCTCGTGTGATGACATTAGCCAGGAGAGAGA 358
      |||:|||||:|||||:||||| :|||:|||||:||||| :||| :||| :|||
122 pThrTrTYrPheserValasnValSeraspalaleuProserSerglUA 139
359 GCCTGGGGCCGACAGCTCCTGTGGGGTCAAGAG.....GACCCC 399
      :: :|||:|||||:|||||:||||| :|||:|||||:||||| :||| :||| :|||
139 spAsPAsPAsPAsPAsPAsPAsPAsPAsPAsPAsPAsPAsPAsPAsP 155
400 GCCAGCCAGCAGTGGGACGACCGCGCTTCACACAGACCTCCAAATGAG 449
      :|||:|||||:|||||:||||| :|||:|||||:||||| :||| :||| :|||
156 ThrLvsProAsnProValAlaProTYrTrpThrSerProGIuInsMetGI 172
450 GCGCCGGGTGATGCGACGCGCGGTGGGTAGTCCGCGGCTCAATGTGG 499
      GCGCCGGGTGATGCGACGCGCGGTGGGTAGTCCGCGGCTCAATGTGG 499
172 uylsLysleuHisAlaValAlaProAlaAlaLysThrValLysPheLysCySP 189
500 TGGCCAGCGGACACCTCGGCGCCGACATCAGTGTGATAAGGACGACAG 549
      :|||:|||||:|||||:||||| :|||:|||||:||||| :||| :||| :|||
189 roserSergLythrProAsnProThrLeuArGIrPleunysasnGIlyS 205
550 GCCTTGACGGGCCAGAG.....GCGCGTAGAGCCAGGAGAA 587
      :|||:|||||:|||||:||||| :|||:|||||:||||| :||| :||| :|||
206 GIUPHe...LysProAsPhtIsarGIleGIyGItyrLysValArGItyrAl 221
588 GAAGTGAGACATGAGCCTGAAGAACCTGCGCGGAGAGACAGCGCAAT 637
      |||:|||||:|||||:||||| :|||:|||||:||||| :||| :||| :|||
221 arThrTrpSerIlIeIleMetAsPserValValProserAsPlysGIyAsnt 238
638 ACACCTGCCGCGGTGTGAACCGCGCGGGGCCCATCAACGCCACTTCAAG 687
      |||:|||||:|||||:||||| :|||:|||||:||||| :||| :||| :|||
238 yrrThytsIIeValGIuaSnGluTyrGIySerIlIeAsnIshTrGIyrGln 254
```



```

48 GlyAspLeuLeuGlnLeuArgCysArgLeuArgAspAspValGlnSerIle 64
162 GACCATGTGGACCAAGATGCC...CGCACCATCCACAGCGGCTGGAGCC 208
64 Asn...TrpLeuArgAspGlyValGlnLeuValGlnSerIle...AsnA 78
209 GCTTCGGGCTGCTGGCGAGGGCTGMAAGTGAACAGTGAAGCGGGAG 258
78 rghArgIleThrGlyGlnGlnValGlnValArgAspSerIleProAla 94
259 GATGGCGGGGTGTACGTGTGCAAGGCCAACAGCGGCTGGAGCGGTGAG 308
95 AspSerGlyLeuThrAlaCysValThrSerSerProSerGlySer...As 110
309 CGTCACTACACCTCTGCTGCTGATGATGATGATGATGATGATGATGAT 358
110 pThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 127
359 GCGTGGGGGGGGGAGAGCTCTGCGGGGCTCAAGAGAGAGAGAGAGAGAG 402
127 sPAspAspAspAspAspAspAspAspAspAspAspAspAspAspAsp 143
403 .....AGCCAGCAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 443
144 ThrLysProAsnArgArgProValAlaProLysTrpThrThrThrThr 160
444 GATGAGGCGCGCGGTGATGCGACGCGCGGTGATGCGGTGATGCGGTGAT 493
160 smetGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLys 177
494 AGTGGGTGGCGCGCGGTGATGCGACGCGCGGTGATGCGGTGATGCGGT 543
177 yScysProSerSerSerGlyThrProAsnProThrLeuArgTrpLeuLys 193
544 GACGAGCGCTTGACGCGCGCGAG.....GCGCGTGGAGCGCGAG 581
194 GlyLysGlnPhe...LysProAsnHisArgIleGlyGlyTrpLysVal 209
582 GAAGAAGAAGTGGACACTGACGCGGTGATGCGGTGATGCGGTGATGCG 631
209 GlyLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 226
632 GCAATATACCTGCGCGGTGATGCGACGCGCGGTGATGCGGTGATGCG 681
226 LysAspLysLysLysLysLysLysLysLysLysLysLysLysLysLys 242
682 TACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 731
243 TyrGlnLeuAspValValGlnLysSerProHisArgProHisProSer 259
732 CACGACCGCGGTGATGCGACGCGGTGATGCGGTGATGCGGTGATGCG 781
259 AsnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 276
782 AGTGAAGTGGCGCGGTGATGCGACGCGGTGATGCGGTGATGCGGTGAT 831
276 eCysLysValLysLysLysLysLysLysLysLysLysLysLysLysLys 292
832 GTGAGGATGCGCGGTGATGCGACGCGGTGATGCGGTGATGCGGTGAT 881
293 IleGlnVal.....AsnGlySerLysLysLysLysLysLysLysLys 304
882 CCAGAAGTTTGTG...GTGCTGCCACGCGGTGATGCGGTGATGCGGTG 928
304 nLeuProGlyValGlnIleLeuLysThrAlaGlyVal.....AsnThr 319
929 ACGGTCTCTACCTCAATAGCTGTATACCGGTGCGCGCGCGCGAGAGAT 978
319 hAspLysGlnLysLysLysLysLysLysLysLysLysLysLysLysLys 335
979 GGGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1028

```

```

336 AlaGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLys 352
1029 CAGCGCGCTTCTCTACCGGTGATGCGGTGATGCGGTGATGCGGTGAT 1072
352 sSerAlaTrpLeuThrValLeuGlnLysLysLysLysLysLysLysLys 366
1073 CACCTGTGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1122
367 .....AlaValMetThrSerProLeuThrLeuGlnLysLysLys 377
1123 ATCGCATCCAGCGCGGTGATGCGGTGATGCGGTGATGCGGTGATGCG 1163
378 IleIleThrCysThrGlyAlaPheLeuIleSerCysMetLeuLysVal 394
1164 GCTCTGTGGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1197
394 IleIleThrCysThrGlyAlaPheLeuIleSerCysMetLeuLysVal 411
1198 .....TGCACCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1242
411 LysMetAlaValHisLysLysLysLysLysLysLysLysLysLysLys 427
1243 GGGAGCGCGCGCGGTGATGCGGTGATGCGGTGATGCGGTGATGCGGT 1275
428 ThrValSerAlaAspSerSerAlaSerMetAsnSerGlyValLeuVal 444
1276 .....CTTCCCTGTGGCGCGGTGATGCGGTGATGCGGTGATGCGGT 1324
444 LysProSerArgLysSerSerSerGlyThrProMetLeuAlaGlyVal 461
1325 AGGAGCATGGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1353
461 eGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLys 470
seq_name: p1r2:A49151
seq_documentation_block:
fibroblast growth factor receptor 1 - Iberian ribbed newt
C:Species: Pleurodeles waltl (Iberian ribbed newt)
C>Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1999
C/Accession: A49151
R:Shl, D.L.; Feige, J.J.; Riou, J.F.; Deslomme, D.W.; Boucaut, J.C.
Development 116, 261-273, 1992
A:Title: Differential expression and regulation of two distinct fibroblast growth fac
A:Reference number: A49151; MUID:93130775
A:Accession: A49151
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-816 <SH1>
A>Note: sequence extracted from NCBI backbone (NCBI:122562, NCBI:122597)
C:Superfamily: basic fibroblast growth factor receptor 1; Immunoglobulin homology: pr
C:Keywords: ATP; growth factor receptor
F:169-230/Domain: immunoglobulin homology <IMM>
F:472-757/Domain: protein kinase homology <KIN>
F:480-488/Region: protein kinase ATP-binding motif
alignment_scores:
Quality: 416.00 Length: 377
Ratio: 1.849 Gaps: 11
Percent Similarity: 59.682 Percent Identity: 28.382
alignment_block:
US-09-598-042A-2 x A49151 ..
Align seg 1/1 to: A49151 from: 1 to: 816
103 GCGCGCGTGGCGCGGTGATGCGGTGATGCGGTGATGCGGTGATGCGGTG 152
44 AlaHisHisGlyAspLeuLeuGlnLeuArgCysArgLeuArgAspAsp 60
153 GCGCGCGTGGCGCGGTGATGCGGTGATGCGGTGATGCGGTGATGCGGTG 202

```

```

60 1HisserleAsn...TropLulysaspGlyValGlnLeu...AlaGlnT 75
203 GAGAGCGCTCCCGCTGCTGCTCCGAGGGCTGAAGTGAAGCGGGAG 252
75 hrasnArgThrArgIleThrGlyAlaGlnValAlaArgAspAlaVal 91
253 CGGAGAGATCCGCGGTGTACGTGTACGAGGCGACCAACGCGCTCCGAG 302
92 GlnGluAspSerGlyLeuValAlaGlySerMetThrHisArgProSerGly 108
303 CCGAGCGTCACTACACCTCGTGTCTGTGATGACATTAGCCAGGGA 352
108 rGlnThrThrPhePheAlaValAlaSerAspArgIleProSerValG 125
353 AGGAGAGCCCGGGCGCCGACAGCTCCTGGGGGTCAAG... 393
125 lnsasparAspAspArgGlnLysSerSerSerGlnGlnValAla 141
394 .....GACCCGCGCAGCAGACACTGGGCGACGCGCGCTT 428
142 GluAsnSerLysProAsnProValAlaProPheThrAla..... 154
429 CACAGAGCCTCCAGATGAGCGCGCGGTGATGCGACGCGCGGTG 478
155 ....HisProGlnLysMetGlnLysLysLeuHisAlaValProAla 170
479 GCTCCGCGCGGTCAAGTGTGCGTGGCGGCGACCCCTCGCGGACATC 528
170 ystThrValLysPheArgCysProAlaGlyGlyThrProSerProThr 186
529 ACSTGATGAGAGCAGCAGCGCTTGACGCGCCAG... 567
187 ArgThrPheLysAsnGlnLysGlnPhe...LysProAspHisArgIleG 202
568 .GCCGCTGAGCCGAGAGAGAGAGTGTGACACTGACCTCAAGACCTGC 616
202 yGlyTyrLysValAlaArgTyrGlnThrProSerIleIleMetAspSer 219
617 GCGCGGAGCAGCGGCAATACACTGCGCGCGGTGAGACCGCGCGG 666
219 alProSerAspLysGlyProTyrThrCysLeuValGlnAsnAsnTyrGly 235
667 GCATCAACGCCACCTCAAGGTGATGATCAGCGGACCGCTTCCA 716
236 SerIleAsnHisThrTyrGlnLeuAspValValGlnLysSerProHis 252
717 GCCCGTCTCAGAGGAGCAGCGCGGTGAGACGCGGTGACTCGGG 766
252 gProIleLeuGlnAlaGlyLeuProAlaAsnGlnThrValProValGly 269
767 GGACCGACGCTTCCAGTGCAGAGGTGCGCAGCGAGCTGAAGCGGTG 816
269 erAsnValAspPheValLysLysValLysSerAspProGlnProHisIle 285
817 CAGTGTGCTGAGCGCGCTG.....GAGTACGCGCGCGAGG 851
286 GlnThrPheLysHisValThrValAsnGlySerLysTyrGlySerAsp 302
852 CCGCGCAACTCCACATCGATGTGGCGCGCAGAACTTTGTGTGCTGC 901
302 Y.....LeuP 304
902 CCACGCGGTGACGTGTGCGCGCGCGAGCGCTCTACCTCAATAG... 948
304 rGluValAlaGlnValLeuValAlaAlaGlyValAlaAsnThrThrAsp 320
949 .....CTGTCAATCACCGCTGCGCGCGCAGAGCAGTGGCGGATG 989
321 MetGluValAlaLeuHisLeuArgAsnValSerPheGlnAspArgLys 337
990 CATGTGCTTGCGCGCAGACACATGGGTACAGTCCGACGCGCGCTCC 1039
337 rThrCysLeuAlaGlyAsnSerIleGlyIleSerHisHisSerIleTrp 354

```

```

1040 TCACGCTGCTGCCAGACCCAAACCGCCAGGCGCAGCTGTGCTCTCG 1089
354 eutThrValValGlu.....AlaIleSer 361
1090 TCCTCGGCGCATAGCCTGCGTGGCGCGGTGTCATGCGCATCCAGCGG 1139
362 GluAsnProValIleMetThrSerProLeuTyrLeuGlnIle..... 375
1140 CGCTGTCTCATCCTGGCGACCCCTGCTCTCG 1170
376 .IleIleTyrCysThrGlyAlaPheLeuIle 385
seq_name: p1r2:S51242

```

```

seq_documentation_block:
heparin-binding fibroblast growth factor receptor - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 06-Jun-1995 #sequence_revision 21-Jul-1995 #text_change 16-Jul-1999
C:Accession: S51242; S51634; S48046
R:Ryan, G.; McBride, G.; McKeenan, W.L.
submitted to the EMBL Data Library, June 1993
A:Description: Exon skipping causes alternation of the COOH-terminus and deletion of
A:Reference number: S51242
A:Accession: S51242
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-353 <YAN>
A:Cross-references: EMBL:L19112
R:Savagner, P.
submitted to the EMBL Data Library, July 1994
A:Reference number: S51634
A:Accession: S51634
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 250-296 <SAY>
A:Cross-references: EMBL:Z35137
R:Savagner, P.; Valles, A.M.; Jouanneau, J.; Yamada, K.M.; Thierly, J.P.
Mol. Biol. Cell 5, 851-862, 1994
A:Title: Alternative splicing in fibroblast growth factor receptor 2 is associated w
A:Reference number: S48046; MUID:95102150
A:Accession: S48046
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 250-276, 'Y', 278-295 <SA2>
A:Cross-references: EMBL:Z35137
C:Superfamily: immunoglobulin homology
C:Keywords: growth factor receptor; heparin binding
F:107-168/Domain: immunoglobulin homology <IMM>

```

```

alignment_scores:
Quality: 411.00 Length: 353
Ratio: 1.995 Gaps: 8
Percent Similarity: 58.357 Percent Identity: 31.728

```

```
alignment_block:
US-09-598-042A-2 x S51242 ..

```

```

Align seg 1/1 to: S51242 from: 1 to: 353
202 TGGAGCGCTTCCGCTGCTGCTGCGCAGGGGTGAAGTGAAGCAGTGA 251
28 TrpGlyYArgPheIleCysLeu..... 34
252 GCGGAGAGATGCGCGGTGTACGTGTGCAAGGCGCAGCAACGCTGGGC 300
35 .....ValIleValThrMetAlaThrLeuSerLeuAla 46
301 .....ACCTGAGCGTCACTACACCTCGTGTGCTGATGATGAC 339
46 rGProSerPheSerLeuValGluAspThrThrLeuGlnProGluAspAla 62

```



[illegible]

```

340 LysLysPro 342
seq_name: p1r2:150128
seq_documentation_block:
  fibroblast growth factor receptor - quail
  C:Species: Coturnix coturnix (quail)
  C:date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
  C:Accession: I50128; 542803
  R:Marcelle, C.; Eichmann, A.; Haley, O.; Brant, C.; Le Douarin, N.M.
  Development 120, 683-694, 1994
  A:title: Distinct developmental expression of a new avian fibroblast growth factor
  A:Reference number: 150128; MUID:94215505
  A:Accession: I50128
  A:Status: preliminary; translated from GB/EMBL/DBJ
  A:molecule type: mRNA
  A:Residues: 1-713 <MAR>
  A:Cross-references: EMBL:X76885; NID:9440139; PIDN:CA54213.1; PID:9440140
  C:Genetics:
  A:Gene: FREK
  C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; pr
  C:Keywords: ATP; growth factor receptor
  F:176-247/Domain: immunoglobulin homology <IM>
  F:377-662/Domain: protein kinase homology <KIN>
  F:385-393/Region: protein kinase ATP-binding motif

alignment_scores:
  Quality: 407.00 Length: 323
  Ratio: 2.035 Gaps: 7
  Percent Similarity: 61.920 Percent Identity: 29.721

alignment_block:
US-09-598-042A-2 x I50128 ..

Align seg 1/1 to: I50128 from: 1 to: 713

322 CTCGCTGCTGGATGACATTAGCCAGGAGAGAGAGCCGTGGGCGCA 371
|||||:|||||: |||: |||: |||||
12 LeuLeuAlaValGlyProAlaAlaSerHisArgGlyGluMetCysProAs 28
372 CAGCTCTCTGGGGTCAAGAGAGAGAGCCGCCAGCCAG..... 408
|||||:|||||: |||||: |||||:
28 pSerLeuAlaSerGlyAspAspAspArgLysSerAspArgLysProH 45
409 .....CAGTGGGACAGAGCCGCGCTTC 429
45 IScLysAspArgSerGluGluProValTyrMetHisArgAlaProTyrTrp 61
430 ACACAGACCCCTCCAGATGAGAGCGCGGTGATGCGAGCCCGTGGGTAG 479
|||||:|||||: |||||: |||||: |||||
62 ThrHisProHisArgMetAspLysLysLeuTyrAlaValProAlaGlyAs 78
480 CTTCCGTGGGCTCAAGTGGCTGGCCAGCGGAGACCTCGGCGGACATCA 529
|||||:|||||: |||||: |||||: |||||
78 nThrValLysPheArgCysProAlaSerGlySerProSerProSerIleA 95
530 CGTGGATGAAGACGACGACGCGCTTGACG.....CGCCACGAGGCC 570
|||||:|||||: |||||: |||||: |||||
95 rGlyrPheLysAsnGlyArgGluPheArgGlyGluHisArgIleGlyGly 111
571 GCGAGCCCGCAGAGAGAAAGTGAAGTGAACAGTGAAGCGCTCAAGACCTGGCGGC 620
|||||:|||||: |||||: |||||: |||||
112 IleArgLeuArgHisGlnHisIleTyrSerLeuAlaMetGluSerValAlaPr 128
621 GGAGAGACAGCGCAATATACACCTGCGCTGTGCAACCGCGGAGGCCA 670
|||||:|||||: |||||: |||||: |||||
128 oSerAspArgGlyAsnTyrThrCysLeuValGluAsnAspArgPheGlyArgI 145
671 TCAACGCCACTTAAGAAGTGGATGTGATCCAGCGGAGCCGTTCCAGGCC 720
|||||:|||||: |||||: |||||: |||||
145 lAsrGlyrSerTyrLeuLeuAspValLeuGluAlaArgSerProHisAspPro 161

```

```

721 GTGCTCAGAGGACGACCCCGTGAACAGACGGTGTGACTTCGGGGGAGC 770
162 TLeuGlnAlaGlyLeuProAlaAsnThrThAlaLeuValGlySerAs 178
771 CAGCTCTTCAGTGCAGAGTGGCGGACGACGTAAGCCGGTATCCAGT 820
178 pValGlnPhePheCysValTyrSerAspAlaGlnProHisLeuGlnT 195
821 GGCTGAAGCGCGTGGATGACGGCGGACGCGCCCAACTCCACCATC 870
195 rPheUlyshHisIleGluVal.....AsnGlySerSerTyr 206
871 GATGTGGGGGCCCAAGAATTGTG...GTGCTGCCACGGGTACGCTGTG 917
207 GlyProAspGlyValProTyrValGlnValLeuUlyshThrAlaSpIle.. 222
918 GTGCGCGCCGACGCGCTCTACCTCAATAAGCTGTATCACCCGTGCC 967
223 .....AsnSerSerGluValGlnValLeuTyrLeuArgAsnValT 236
968 GCCAGAGCATGCGGCGATGATCATCTGCTTGGCGCAACACCATGGCC 1017
236 hrMetGlnAspAlaGlyGluTyrThrCysLeuAlaGlyAsnSerIleGly 252
1018 TACAGGTCCGACGCGCTCTCTACCGCTGCGACGACCCAAACCGCC 1067
253 LeuSerTyrGlnSerAlaThrPheUlyshThrValLeuProGlnGluLeuVal 269
1068 AGGCGCACCTGTGGCGCTCTCTCTGCTGCGGACCTGACCTGCGCGCC 1117
269 LHisGlnAlaGlnTyrThrSerGlnAlaUlyshTyrThrAsp.....I 282
1118 TGGTTCATCGGCATCCAGCGCGGCTGTCTCATCTCGGCGACCTGTCT 1167
282 LLeuIleTyrThrSerGlySerLeuAlaValAlaMetAlaLeuIleTyr 298
1168 CTGTGCTTTCAGGAGCCAGAAAGCCGTGCACCCCGCGCTGCGCC 1217
299 ValValLeuUlyshArgMetGlnThrGln...SerSerUlyshProLeuGln 314
1218 TCCCTGTGCTGGCGACCGC 1236
314 uProMetAlaValAlaHisUlysh 320

```

seq\_name: p1r2:PN0020

seq\_documentation\_block:

fibroblast growth factor receptor 2 - mouse  
C:Species: Mus musculus (house mouse)

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 05-Nov-1999  
C:Accession: P00020:148348

R:Faasel, N.J.; Bernard, M.; Deglon, N.; Rousseaux, M.; Eisenberg, R.J.; Bron, C.; Cohen, Blochem. Biophys. Res. Commun. 178, 8-15, 1991

A:Title: Isolation from mouse fibroblasts of a cDNA encoding a new form of the fibroblast A:Reference number: P00020; MUID:91298988

A:Accession: P00020

A:Molecule type: mRNA

A:Residues: 1-361 <FAS>

R:Deglon, N.; Krapp, A.; Bron, C.; Faasel, N. Blochem. Biophys. Res. Commun. 174, 1337-1342, 1991

A:Title: Translocation of the yeast dolichol-phosphate-mannose synthase into microsomal A:Reference number: 148348; MUID:91144617

A:Accession: 148348

A:Molecule type: mRNA

A:Residues: 1-360 <RES>

A:Cross-References: EMBL:X57790; NID:9499309; PIDN:CAA0932.1; PID:9499310

C:Genetics:

A:Gene: MB-2

C:Superfamily: Immunoglobulin homology

C:Keywords: growth factor receptor

F:169-230/Domain: Immunoglobulin homology <IMM>

alignment\_scores:

Quality: 390.50 Length: 320  
Ratio: 1.842 Gaps: 10  
Percent Similarity: 66.250 Percent Identity: 30.312

alignment\_block:

US-09-598-042a-2 x PN0020

Align seg 1/1 to: PN0020 from: 1 to: 361

```

112 GCGCGCATATGCGGCTGATGTCAGTCCAGTGAAGGAGGACCGCCCGCT 161
48 GLYAspLeuGlnLeuArgCysArgLeuArgAspAspValGlnSerIle 64
162 GACCATGTGGACCAAGATGCG...CGCACATCCACACCGCGCTGAGCC 208
64 eAsn...TrpLeuArgAspGlyValGlnLeuValGlnSer.....AsnA 78
209 GCTCCGCGTGTGCGCGGCGGCGTGAAGTGAAGAGGTGAGCGGAG 258
78 rGThrArgIleThrGlyGlnGluValGlnValArgAspSerIleProAla 94
259 GATGCGGCGGTGTGATGTGCAAGCGCACCAAGCGCTTGCAGCCTGAG 308
95 AspSerGlyLeuTyrAlaCysValThrSerSerProSerGlySer...As 110
309 CGTCAACTACACCTGCTGCTGTGATGATCATTTAGCCAGGAGGAGAGA 358
110 pThrThrTyrPheSerValAsnValSerAspAlaLeuProSerSerGln 127
359 GCCTGGCGCGCGACAGCTCTCTGCGGCTCAAGAG.....GACCC 399
127 spAspAspAspAspAspAspAspSerSerSerGlnUlyshGlnThrAspAsn 143
400 GCCAGCGCATGTGGCGACGCGCGCTTGCACACACCGCTTGCAGATGAG 449
144 ThrUlyshProAlaProValAlaProTyrThrThrSerProGlnUlyshMet 160
450 GCGCGGCGGTGATGCGACGCGCGGTGATGCTGCGGCTCAATGCG 499
160 UlyshUlyshLeuHisAlaValProAlaAlaUlyshThrValUlyshPheUlyshCysP 177
500 TGCGCGCGCGCGACCGCTGCGCGCGCGACATCATGATGATGAAGACGACG 549
177 roSerSerGlyThrProAsnProThrLeuArgTyrLeuUlyshAsnGlyUlysh 193
550 GCCTTACCGCGCGCGAG.....GCCGTGACCGCGAGAGAA 587
194 GluPhe...LysProAspHisArgIleGlyGlyTyrUlyshValAlaArgTyrAl 209
588 GAAGTGACATGAGCGCTGAGAACCTGCGCGCGAGGAGGACGCGCGAAT 637
209 aThrThrSerIleIleMetAspSerValAlaProSerAspUlyshGlyAsnT 226
638 ACACTGCGCGCGTGTGAACCGCGCGCGCGCATCAACGCGCATCAAG 687
226 yThrCysIleValGlnUlyshGlyUlyshSerIleAsnHisThrTyrGln 242
688 GTGATGTGATCCAGCGGAGCGCGCTTCCAGCGCGCTGCTCAGACGCGCA 737
243 LeuAspValValIleUlyshSerProHisAspGlyProHisIleProHisIleUly 259
738 CCGCGTGAACAGCAGCGTGTGCGGAGGAGGACGCGCGCTTCTTCAGTGA 787
259 uProAlaAsnUlyshThrValAlaLeuUlyshSerAsnValGluPheMetCys 276
788 AGGTGCGCGACGCGATGAGCCGCGTATCCAGTGTGAGCGCGTGGAG 837
276 yValTyrSerAspProGlnProHisIleGlnThrPheUlyshHisIleGln 292
838 TACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 867

```

```

293 Val.....:AsnGlySerIysIleGlyProAspAsnLeuPr 304
888 GTTGTG...GTGCTGCCACAGGTTGACGTGTGCTGGCGCCACAGGCT 934
304 CTYValGlnIleLeuIysThrAlaGlyVal.....AsnThrIAspL 319
935 CCTACCCCAATAAAGTGTCTCATCAACCGTGGCCGCCACAGAGATGGCGG 984
319 YsgIuMeGluValIleuHsIleuArgAsnValSerPheGluAspAlaGly 335
985 ATGTACATCTGCTTGGCGCCACACACATGGCTACAGCTTCCGACGCGC 1034
336 GluIyrThrCysIleuAlaGlyAsnSerIleAlaIleuGluAlaGlyPro 352
1035 CTTCCTCAC 1044
352 aValMethr 355

```

seq\_name: pir2:A35969

```

seq_documentation_block:
heparin-binding growth factor receptor K-sam precursor - human
C:Species: Homo sapiens (man)
C>Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 16-Jul-1999
C:Accession: A35969
R:Factor1_Y.: Odaagiri, H.; Nakatani, H.; Miyagawa, K.; Naito, K.; Sakamoto, H.; Katoh,
Proc. Natl. Acad. Sci. U.S.A. 87, 5983-5987, 1990
A:Title: K-sam, an amplified gene in stomach cancer, is a member of the heparin-binding
A:Reference number: A35969; MUID:90332706
A:Accession: A35969
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-682 <HAT>
A:Cross-references: GB:M35718; NID:q186777; PIDN:AAA36152.1; PID:q186778
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
C:Keywords: ATP; autophosphorylation; growth factor receptor; heparin binding; phosphop
F:83-144/Domain: immunoglobulin homology <IM>
F:389-674/Domain: protein kinase homology <KIN>
F:397-405/Region: protein kinase ATP-binding motif

```

alignment\_scores:

```

Quality: 390.50 Length: 354
Ratio: 1.905 Gaps: 9
Percent Similarity: 57.910 Percent Identity: 30.791

```

alignment\_block:

US-09-598-042a-2 x A35969 ..

Align seg 1/1 to: A35969 from: 1 to: 682

```

202 TGGAGCGCTTCCGCGTGTGCTGCCAGAGGGCTGAAGTGAGAGAGTGGA 251
4 TrpIlyArgPheIleCysLeu..... 10
252 GCGGAGAGATCCGCGGTGTACGTGTGCAGACCAACAGCTTCGCGC. 300
11 .....ValValValThrMetAlaThrLeuSerLeuAla 22
301 .....AGCCTGAGCTCAACTACACCTTCGCTGCTGTGATGAC 339
22 rgProSerPheSerLeuValGluAspThrThrLeuGluProGluAspAla 38
340 ATTAGCCAGGAGAGAGAGCTGGGCGCCAGACGCTCTCTGGGGGTCA 389
39 IleSerSerGlyAspAsp.....GluAspAspThrAspGlyAl 51
390 AGAGAGCCCGCCAGCCAG.....CACTGGGAGAGAGAGCCGCGCTTCA 430
51 aGluAspPheValSerGluAsnSerAsnAsnLysArgAlaProIyrTrp 68
431 CACAGCCCTCAAGATGAGGCGCGGTGATGCACAGGCCCGCTGGGTAC 480

```

```

68 hrasnThrGluIysMetCyluIysArgLeuHsIAlaValProAlaIalaasn 84
481 TCCGTGCGGCTCAAGTGTGCGGACAGCGGACCCCTGGCCGACATCAC 530
85 ThrValIysPheArgCysProAlaGlyGlyAsnProMetProThrMetAr 101
531 GTGATGAGAGACGACGACGCGCTTGACG.....CGCCACAGAGCGG 571
101 gTrpLeuIysasnGlyLysGluPheLysGlnGluHsIArgIleGlyLyr 118
572 CTGAGCCAGAGAGAAGAAGTGGACACTGACCTGACAGACCTGGAGCGG 621
118 yTrIysValArgAsnGlnHsItrPserIleuIleMetGlnSerValValPro 134
622 GAGACAGCGGCAATATACCTGCGCGCTGTGCAACCGCGGCGGCAT 671
135 SerAspLysGlyAsnTrpThrCysValValGluAsnGluTrpGlySerI 151
672 CAAGCCCACTACAAAGTGTGATGTGATCCAGCGGACCGCTTCCAAGCCG 721
151 easnHsIThrTrpHsIleAspValValGluArgSerProHsIArgProI 168
722 TGCTACAGAGCGACGACCCCGTGAACAGACGCGTGTGCGGAGGAC 771
168 IeLeuGlnAlaGlyLeuProAlaAsnAlaSerThrValValGlyLysp 184
772 AGTCCTTCCAGTGCAGAGTGCAGCGACGAGCTGAGACCGGTGATCCATG 821
185 ValGluPheValCysLysValTyrSerAspAlaGlnProHsIleGlnTr 201
822 GCTGAGCGCGTGGAG.....TAGGCGCCAGAGGCGGCC 856
201 pIleLysHsIValGluLysasnGlySerLysTrpGlyProAspGly.... 216
857 ACAACTCCACCATGATGTGGGCGCGCAGAGATTGTGTGCTGGCCACAG 906
217 .....LeuProIyr 219
907 GGTGACGTG.....TGTCGCGGCGCGACGCGCTCTACCTCAATTAAGCT 950
220 LeuIysValLeuLysHsIserGlyIleAsnSerSerAsnAlaGluVal 236
951 GCTCATCACCGTGCAGGACGATGCGGACATGTACATCTGCTTG 1000
236 uAlaLeuPheAsnValThrGluAlaAspAlaGlyLyrIleCysLysV 253
1001 GCGCCAAACACATGGGCTACAGCTTCCGACGCGCTTCTCCTCAGCTGTG 1050
253 aIserAsnTrpIleGlyGlnAlaAsnGlnSerAlaTrpLeuThrValLeu 269
1051 CCAGACCCAAACCGCCAGGCGCACCTGTGCTCTCCTGCGGCGCAC 1100
270 ProLysGlnGlnAlaProGly.....ArgGluLysGlnIleTh 282
1101 TAGCCTGCGTGGCGCGGTGTATCGCATCCACAGCGCGCGCTTCTTCA 1150
282 rAlaSerProAspTrpLeuGlnIleAlaIleIleTrpCysIleGlyAlaPheL 299
1151 TCCTGGGACCGCTGCTCTGTGCTTTCAGAGCCCAAGAAGCGCGTC 1200
299 euIleAlaCysMetValValThrValIleLeuCysArgMetLysasnThr 315
1201 ACCCCCGCGCT 1212
316 ThrIysLysPro 319

```

seq\_name: pir2:A54846

```

seq_documentation_block:
fibroblast growth factor receptor a precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 28-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 16-Jul-1999

```



```

21  aargProserPheSerLeuVal..... 28
356 AGAGCGTGGGGCCGACAGCTCTGCGGGGTCAGAGAGACCCCGCCAGC 405
29 ..... 33
406 CAGCAGTGGGCGACGCGCTTTCACACAGCCCTCCAGATAGAGCGCG 455
34 GUPROGUGLUGLALAPROTYTTRPThrAsnThrGluLysMetGluLysAr 50
456 GGTATTCGACGCGCGCTGGTAGCTCCGTGGGCTCAAGTCCGTGGCCA 505
50 GLeuHISAlaValProAlaAlaAsnThrValLysPheArgCysProAlaG 67
506 GCGGCGACCCCTGGCGCCGACATCAGCTGGATAGAGAGACAGCGCTTG 555
67 LylAsnProThrProThrMetArgTrpLeuLysAsnLylLysGluPhe 83
556 AGC.....CGCCAGAGCGCGCTGAGCCCGAGAGAGAAAGATGGAGC 596
84 LysGlnGlnHISArgLlLeGlyLylLysValArgAsnGlnHISTrpSe 100
557 ACTGAGCTGAAGAACTGCGCGCGGAGACAGCGCAAAATACACTGGCC 646
100 rLeuLleMetGluSerValValProSerAspLysGlyAsnGlyTrpCysL 117
647 GCGTTCGACGCGCGCGCGCCATCAACGCGCAACTACAGTGGATGTG 696
117 euValGlnAsnGlnLylLysTrlLeAsnHISThrLysHISLeuAspVal 133
697 ATCCAGCGGACCGCTTCAGAGCGCGGTCCAGACAGCGCCCGCTGAA 746
134 ValGlnLylSerProHISArgProLleLeuGlnAlaGlyLeuProAlaAs 150
747 CAGCAGGTGGACTTGGGCGGAGCAGCGCTCTCCAGTCAAGTGGCGCA 796
150 nAlaSerThrValValGlyLylAspValGluPheValCysLysValLys 167
797 GCGAGCTGAAGCGCGTATCCAGTGGCTGAACGCGCTGGAG..... 837
167 eArspAlaGlnProHISArgLlLeGlnTrlLeLysHISValGlnLysAsnGly 183
838 .....TACGGCGCGGAGCGCGCCACAACCTCCAGATCGATGGCGG 881
184 SerLysTrlLylLysProAspLyl..... 190
882 CCAGAGTTTGTGTGCTGCCCGCAGCGGTGACGTG.....TGGTCGCGGC 925
191 .....LeuProTrlLleuLysValLeuLysHISSerGlyI 202
926 CCGACGCTCTACCTCAATAAGCTGTATCCAGCGCGCGCGCCAGAGC 975
202 LeAsnSerSerAsnAlaGlnValLeuAlaLeuPheAsnValThrGluMet 218
976 GATGGGGGATATACATCTGCTGGCGCCACACACATGGGTAAAGCTT 1025
219 AspAlaGlnGlyLylLysValSerAsnTrlLeGlyGlnAlaAs 235
1026 CCGCAGCGCTCTCTACCGCTGCTGCCAGACCCAAACCGCAGCGCGCAC 1075
235 nGlnSerLlArTrpLeuThrValLeuProLysGlnAlaPro..... 249
1076 CTGTGGCTCTCTGCTGCGCGCACTAGCGCTGCCGCGCGCTGTATC 1125
250 .....ValArgGlnLylLysGlnLlLeThrAlaSerProAspTrlLeuLlLe 264
1126 GGCATCCAGCGCGCGCTGTATCCGCGGCGCGCGCGCTCTCTGTGGT 1175
265 AlAlleTrlLysLlLeGlyValPheLeuLlLeAlaCysMetValValThrVa 281
1176 TTGCCAGCGCGCAGAGAGCGCTGCACCCCGCGCT 1212
281 lLlPheCysArgMetLysThrThrLysLysPro 293

```

```

seq_name: p1r2:S51635
seq_documentation_block:
  fibroblast growth factor receptor 2b, keratinocyte growth factor receptor - rat
  C/Species: Rattus norvegicus (Moriway rat)
  C/Date: 15-Jul-1995 #sequence revision 21-Jul-1995 #text_change 16-Jul-1999
  C/Accession: S51635; S51636; S48048; S48047
  R/Savagner, P.
  submitted to the EMBL Data Library, July 1994
  A/Reference number: S51634
  A/Accession: S51635
  A/Status: preliminary
  A/Molecule type: mRNA
  A/Residues: 1-705 <SAV>
  A/Cross-references: EMBL:Z35138; NID:9551271; PIDN:CAA84510.1; PID:9551272
  A/Accession: S51636
  A/Status: preliminary
  A/Molecule type: mRNA
  A/Residues: 37-172,'T',174-314,'TV',315-340,'T',342-705 <SA3>
  A/Cross-references: EMBL:Z35139; NID:9551273; PIDN:CAA84511.1; PID:9551274
  R/Savagner, P.; Valles, A.M.; Jouanneau, J.; Yamada, K.M.; Thliery, J.P.
  Mol. Biol. Cell 5, 851-862, 1994
  A/Title: Alternative splicing in fibroblast growth factor receptor 2 is associated w
  A/Reference number: S48046; MUID:95102150
  A/Accession: S48048
  A/Status: preliminary; nucleic acid sequence not shown
  A/Molecule type: mRNA
  A/Residues: 37-135,'R',137-314,'TV',315-340,'T',342-390,'K',392-705 <SA2>
  A/Cross-references: EMBL:Z35139
  A/Accession: S48047
  A/Status: preliminary; nucleic acid sequence not shown
  A/Molecule type: mRNA
  A/Residues: 1-135,'R',137-340,'T',342-390,'K',392-705 <SA4>
  A/Cross-references: EMBL:Z35138
  C/Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; pr
  F/57-118/Domain: immunoglobulin homology <IM>
  F/363-648/Domain: protein kinase homology <KIN>
  F/371-379/Region: protein kinase ATP-binding motif

alignment_scores:
  Quality: 369.50      Length: 284
  Ratio: 2.099        Gaps: 5
  Percent Similarity: 61.972      Percent Identity: 31.690

alignment_block:
US-09-598-042a-2 x S51635 ..

Align seg 1/1 to: S51635 from: 1 to: 705
391 GAGAGACCCCGCAGCGAGTGGGACAGCGCGCTTCACAGCGCTC 440
||||| : : : : : : : : : : : : : : : : : : : : : : : :
29 GluAspTrlThrLeuGlnProGluLylAlaProTrlLysThrAsnThrG 45
441 CAAGATGAGCGCGCGGTATGSCAGCGCGCGCTGGTAGCTCCGTGGCG 490
: : : : : : : : : : : : : : : : : : : : : : : :
45 ulysMetGluLysArgLeuHISAlaValProAlaAlaAsnThrValLysP 62
491 TCAATGCGCTGGCGCAGCGGCGCACTCCGCGCGCAATCACTGATGAAG 540
: : : : : : : : : : : : : : : : : : : : : : : :
62 heArgCysProAlaGlnLylLysAsnProThrProThrMetArgTrpLeuLys 78
541 GAGCAGCAGCGCTTGAGC.....CGCCAGAGCGCGCTGAGCCGAG 581
: : : : : : : : : : : : : : : : : : : : : : : :
79 AsnGlnLylGlnPheLysGlnLlHISArgLlLeGlyLylLysValAr 95
582 GAAGAAGAGTGGACACTGAGCTGAAGAACTGCGCGCGGAGAGACAGCG 631
: : : : : : : : : : : : : : : : : : : : : : : :
95 gAsnGlnHISTrpSerLeuLleMetGluSerValValProSerAspLysG 112
632 GCAATATACCTTCGCGCGCTGTGCAACGCGCGCGCGCCATCAACCGCAC 681

```

```

112 LysnYrThrcYsLeuValGluAsnGluYrGlySerIleAsnHstHr 128
129 TyrHisLeuAspValValArgGlySerThrGlnArgProIleuGlnG 145
732 CAGGACCCCGTGAACAGAGCGGTGACTGGGGGACCACTCTCC 781
145 yGlyLeuProAlaAsnGlySerThrValValGlyAspValGluPhe 162
782 AGTGAAGTGGCCAGCAGTGAAGCGGTGATCCAGTGGTGAAGCC 831
162 alystysValYrSerAspAlaGlnProHisIleGlnTrpIleYsHis 178
832 GTGAG:.....TACGGCCGAGGCGCCGACCAACTCCAC 866
179 ValGluYsAsnGlySerYrGlyProAspGly..... 190
867 CATCGATGTGGGGCCAGAAATTGTGTGTCGCCACGCGTGAAGTG 915
191 .....LeuProYrLeuYsVal 197
916 ....TGTGCGCGCCGAGCGCTCTACCTCAATAAGCTGTCATCACC 960
197 eulysHisSerGlyIleAsnSerSerAsnAlaGluValLeuAlaLeuPhe 213
961 CGTGGCCGCGCAGAGCATGCGGCGATGATCATGCTGCTGGCCCAAC 1010
214 AsnValThrIleMetAspAlaGlyGluYrIleCysYsValSerAsnY 230
1011 CAGGGGTACAGCTTCGCGAGCGCTTCCTCAGCGTGCAGACCCAA 1060
230 rIleGlyGlnAlaAsnGlnSerAlaTrpLeuThrValLeuProYsGln 247
1061 AACCGCCAGGCGCCACCTGTGCGCTCTCTCTCGGCGCACTAGCTCCG 1110
247 lnaIlePro.....ValArgIuYsGluIleThrAlaSerPro 259
1111 TGGCCCTGTGTCATCGGCATCCAGCGCGGCTGTCTTCATCTCGGCG 1160
260 AspYrLeuGlnIleAlaIleYrCysIleGlyValPheLeuIleAla 276
1161 CCGCTCTCTGCGGCTTGGCAGGCGCCAGAAAGCGCTGCACCGCCGCG 1210
276 sMeValValThrValIlePheCysArgMetYsThrThrYsLysP 293
1211 CT 1212
293 ro 293
seq_name: p1r2:149293
seq_documentation_block:
fibroblast growth factor receptor-1, short isoform precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 16-Jul-1999
C:Accession: I49293
R:Jin, Y.; Pasumarthi, K.B.; Bock, M.E.; Lytras, A.; Karamdi, E.; Cattlin, P.A.
J. Mol. Cell. Cardiol. 26, 1449-1459, 1994
A:Title: Cloning and expression of fibroblast growth factor receptor-1 isoforms in the m
A:Reference number: I49289; MUID:95205422
A:Accession: I49293
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-733 <RES>
A:Cross-references: EMBL:U23445; NID:q733537; PIDN:AA52183.1; PID:q733538
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
C:Keywords: ATP; growth factor receptor
F:181-254/Domain: immunoglobulin homology <IM>
F:387-672/Domain: protein kinase homology <KIN>
F:395-403/Region: protein kinase ATP-binding motif

```

```

alignment_scores:
Quality: 365.00 Length: 367
Ratio: 1.659 Gaps: 11
Percent similarity: 59.946 Percent identity: 28.610
alignment_block:
US-09-598-042a-2 x I49293
Align seg 1/1 to: I49293 from: 1 to: 733
337 GACATTAGCCAGGAGAGAGAGAGCTGGGGCCGACAGCTCTGGGG 386
111 .....
31 AspAlaLeuProSerSerGlnAspAspAspAspAspAspSerSerSe 47
387 TCAGAGGACCCCGC.....AGCCAGCAGTGGGACGAC 421
47 rGluGluYsGluThrAspAsnThrLysProAsnArgAlaGluValAla 64
422 CGCGCTTCAACAGCCCTCCAGATGAGCGCGCGGTGATCCAGCGCC 471
64 rGlyTrpThrSerProGluYsMetGluYsIleuHisAlaValPro 80
472 GTGGGTAGCTCCGTGGGCTCAAGTGCAGCGGCGGCGACCTCGGCC 521
81 AlaAlaYsThrValYsPheYsCysProSerSerGlyThrProAsnPr 97
522 CGACATCAGCTGATAGAGACGACGAGCGCTTGACGCCGCCAGAG... 567
97 oThrLeuAlaYrPheYsAsnGlyYsGluPhe...LysProAsnHisA 113
568 .....GCCGCTGAGCCAGAGAGAGAGAGAGTGAACACTGAGCCGAG 609
113 rGluIleGlyYrYsValArgYrAlaThrTrpSerIleIleMetAsp 129
610 AACCTGCGCGCGGAGAGAGAGCGCAATACCTGCGCGGTGCAACCG 659
130 SerValValProSerAspYsGlyAsnYrThrCysIleValGluAsnG 146
660 CGGCGCGCCATCAACGCCACCTACAGAGGTGATGATCCAGCGGAC 709
146 uYrGlySerIleAsnHisThrYrGlnLeuAspValValGluArgSerP 163
710 GTTCCAGGCGCGTGTCCAGAGCGACCGCCGCTGAACAGAGGTGAG 759
163 roHisArgProIleLeuGlnAlaGlyLeuProAlaAsnGluThrValAla 179
760 TTCGGGGGACACAGTCTTCCAGTGCAGAGTGGCAGCGAGCTGAAGC 809
180 LeuGlySerAsnValGluPheMetCysYsValYrSerAspProGlnPr 196
810 GTGATCCAGTGGCTGAGCGCGGTGAGTACGCGCGCGCGCGCAC 859
196 oHisIleGlnTrpLeuYsHisIleGluVal.....Asn 208
860 ACTCCACCATCATGTGGCGCGCCAGAAATTGTG...GTGTCGCCACG 906
208 lYserYsIleGlyProAspAsnLeuProYrValGlnIleLeuYsThr 224
907 GTGACGCTGTGTCGGCGCGCGGCTCTTACCTCAATAGCTGTCAT 956
225 AlaGlyVal.....AsnThrThrAspYsGluMetGluValLeuHisSe 239
957 CACCGCGCGCGCGCAGAGCATGCGGCGCATGTCATCTGCGCTGGCGCA 1006
239 uArgAsnValSerPheGluAspAlaGlyGluYrThrCysLeuAlaGlyA 256
1007 ACACCATGGGCTCAGCTTCCAGCGCGCTTCTTCAACCGTGTGCCA... 1053
256 snSerIleGlyLeuSerHisHisSerAlaTrpLeuThrValLeuGluAla 272
1054 ...GACCCAAACCGCGAGGCGCACCTGTGGCTCTCTCTGGCGCAC 1100
::: :::::

```

```

273 LeuGIuGIuArgPro.....AlaValMetH 281
1101 TAGCTGCGCGCGCGTGCATCGCATCCAGCGCGCGCTGTTC 1150
      |||||      :|||      :|||      :|||
281 rSerProLeuYrLeuGIuLeuIleIleIleYrCysThrCIyAlaPheLeuI 298
1151 TC.....CTGGGACCGCTGCTCCGTGGCTTGGCCAGGCCAGAG 1101
      |||||      :|||      :|||      :|||
298 lSerCysMetLeuGIuSerValIleIleIleYrLysMetLysSerGIyThr 314
1192 AAGCCG.....TGACACCGCGCGCTGCCCTCC 1220
      |||||      :|||      :|||      :|||
315 LysLysSerAlaPheHISserGIuMetAlaValHISLysLeuAlaLysSe 331
1221 CTTCCCTGGGACCGCGCGCGCGCGCGCGCGCGCGAGACA 1270
      :|||      :|||      :|||      :|||
331 rIleProLeuArgArgInValThrValSerAlaAspSerSerAlaSerM 348
1271 AGGAC.....CTTCCTGTGGCGCGCGCGCTCAGCGCT 1302
      :|||      :|||      :|||      :|||
348 etaASerGIyValLeuLeuValArgProSerArgLeuSerSerSerGIy 364
1303 GGCCCTGTGGCGCTGTGTAGAGCATGGTCTCCGCGAGCGCCCA 1352
      |||||      :|||      :|||      :|||
365 ThrProMetProAlaGIyValSerGIuYrGIuLeuProGIuAspProAr 381
1353 G 1353
381 g 381

```

```

seq_name: p1r1:JC1450
seq_documentation_block:
Fibroblast growth factor receptor 4 - rat
N.Contains: protein-tyrosine kinase (EC 2.7.1.112)
C.Species: Rattus norvegicus (Norway rat)
C.Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 01-Dec-2000
C.Accession: J01450, P01091
R.Horllick, R.A. Stack, S.L.; Cooke, G.M.
Gene 120, 291-295, 1992
A.Title: Cloning, expression and tissue distribution of the gene encoding rat fibroblast
A.Reference number: J01450; MUID:93013049
A.Accession: J01450
A.Molecule type: mRNA
A.Residues: 1-650 <HOR>
A.Cross-references: GB:M91599; NID:9204137; PIDN:AAA41157.1; PID:9204138
R.Lai, C.; Lemke, G.
Neuron 6, 691-704, 1991
A.Title: An extended family of protein-tyrosine kinase genes differentially expressed in
A.Reference number: P01083; MUID:91222560
A.Accession: P01091
A.Molecule type: mRNA
A.Residues: 465-518 <LAI>
A.Experimental source: sciatic nerve
C.Genetics:
A.Gene: bGFRA; tyro-9
C.Function:
A.Description: receptor mediating effects of fibroblast growth factor
A.Note: expressed in normal lung; expressed in some carcinomas
C.Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
C.Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;
F:11-72/Domain: immunoglobulin homology <IM1>
F:110-181/Domain: immunoglobulin homology <IM2>
F:118-238/Domain: transmembrane #status predicted <TM>
F:313-650/Domain: intracellular #status predicted <INT>
F:313-598/Domain: protein kinase homology <KIN>
F:321-329/Region: protein kinase ATP-binding motif
F:104, 136, 157, 168/Binding site: carboxylate (Asn) (covalent) #status predicted
F:351, 368, 460/Active site: Lys, Glu, Asp #status predicted
F:465, 478/Binding site: magnesium (Asn, Asp) #status predicted
F:491/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

```

```

alignment_scores:
Quality: 364.00      Length: 332
Ratio: 1.867      Gaps: 13
Percent Similarity: 58.735      Percent Identity: 31.024
alignment_block:
US-09-598-042A-2 x JC1450
Align seg 1/1 to: JC1450 from: 1 to: 650
445 ATGAGGCGCGCGGTGATCGACGCGCGCGGTGAGTCCGTGCGCTCAA 494
      |||||      :|||      :|||      :|||
1 MetGIuLysLysLeuHISAlaValProAlaGIyAsnThrValLysPheAr 17
495 GTGGGTGGCCAGCGGACCGCTCGCGCGCGACATCAGCTGATGAAGACG 544
      :|||      :|||      :|||      :|||
17 gCysProAlaAlaGIyAsnProMetProThrIleHISrLeuLysAsnG 34
545 ACCAGGCTTG.....ACGGCCAGAGCGCGCTGAGCCAGAGAG 585
      |||||      :|||      :|||      :|||
34 LysAlaAlaPheHISGIyLysAsnArgIleGIyLysArgLeuAlaHIS 50
586 AAGAAGTGCACATGAGCTGAGACCTGAGACCTGCGCGCGAGACAGCGCA 635
      :|||      :|||      :|||      :|||
51 GlnHISrPserLeuValMetGIuSerValAlaProSerAspArgGIyTh 67
636 ATACACCTGCGCGCGTGTGAAACCGCGCGCGCGCATCAACGCCACCTACA 685
      :|||      :|||      :|||      :|||
67 rTyrThrCysLeuValGIuAsnSerLeuGIySerIleArgYrSerYrL 84
686 AGGTGATGTATCCAGCGGACCGCTTCCAGCGCGCTGTCTACAGGACG 735
      :|||      :|||      :|||      :|||
84 eulAspValLeuGIuArgYrSerProHISArgProIleLeuGIuAlaGIy 100
736 CACCCCGTGAACAGCAGCGTGTGACTCGGCGGAGACAGCTCTTCCAGTG 785
      :|||      :|||      :|||      :|||
101 LeuProAlaAsnThrThrAlaValAlaGIySerAsnValGIuLeuLeu 117
786 CAAGTGGCGGACGACGTGAAGCGGTGATCCAGTGTGAGCGCGCTGG 835
      :|||      :|||      :|||      :|||
117 slyValValYrSerAspAlaGIuProHISIleGIuThrLeuLysHISIle 134
836 AGTACGCGCGCGAGCGCGCGCACACATCCACATCATGATGTGGCGGCG 885
      :|||      :|||      :|||      :|||
134 alile.....AsnGIySerSerLeuGIuAlaAspGIyPhe 145
886 AAGTTTGG.....GTGTCGCCAGCGGTGACGTGTGTCGCGCGCGAGCG 932
      :|||      :|||      :|||      :|||
146 ProTyrValGIuValLeuLysThrThrAspIle.....AsnSe 158
933 CTCCTACCTCAATAAGCTGTATCACCCGTGCCCGCAGAGACGATGCG 982
      :|||      :|||      :|||      :|||
158 rSerGIuValGIuValLeuYrLeuArgAsnValSerAlaGIuAspAlaG 175
983 GCATGTACATCTGCTTGGCGCGCAACACATGGCTACAGCTTCCGACG 1032
      :|||      :|||      :|||      :|||
175 LysGIuYrThrCysLeuAlaGIyAsnSerIleGIyLeuSerTyrGIuSer 191
1033 GCTTTCCTCACCGGTGTCGCGAGACCCAAAACCGCGAGGCGACCTGTGGC 1082
      :|||      :|||      :|||      :|||
192 AlaThrPheThrValLeuPro..... 198
1083 CTCCTGTCTCTGGGCACTAGCTGCGCTGG..... 1113
      :|||      :|||      :|||      :|||
199 .....AlaGIuGIuGIuAspLeuAlaTrpThrThrAlaThrSerGIuA 213
1114 .....CCGTGTATCGGCATCCAGCGCGCGCTTTCATC 1152
      :|||      :|||      :|||      :|||
213 laArgTyrThrAspIleLeuLeuYrValHISGIySerLeuAlaLeuVal 229
1153 CTGGGACCGCTGTGCTGTGGCGCGCGAGGCGAAGAAGCGGTGCAC 1202
      :|||      :|||      :|||      :|||
230 LeuLeuLeuLeuAlaGIyValTyrHISArgGln..... 241

```

1203 CCCCCGCGCTGCCCCCTCCCTGCGTGGGACGCGCCGCGGGAGCGGCC 1252  
 242 .....AlaLeHisGlyHisHis.SerArg...GlnPro 251  
 1253 GCGACCGCAGCGGAGACAGACGACCTCCCT..... 1282  
 252 ValThrValGlnIlySerArgPheProLeuAlaArgGlnPheSerLe 268  
 1283 .....CGTTGGCGCGCCCTGACGCGCTGGCCCTGGTGGGCGCTGTG 1322  
 268 uGluSerArgSerSerGlyLysSerSerLeuSerLeuValArgGly...V 284  
 1323 TGAGAGACATGGGCTCTCGGACAGCCCGCCAGCACTTACTGGG 1364  
 284 aAlaArgLeuSerSer.SerGly...ProProLeuLeuThrGly 296

seq\_name: p1r2:A56795

seq\_documentation\_block:  
 fibroblast growth factor receptor 1 beta-isoform - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 03-Nov-1995 #sequence\_revision 03-Nov-1995 #text\_change 16-Jul-1999  
 C:Accession: A56795  
 R:Kim, E.G.; Kwon, H.M.; Burrow, C.R.; Ballermann, B.J.  
 Am. J. Physiol. 264, F66-F73, 1993  
 A:Title: Expression of rat fibroblast growth factor receptor 1 as three splicing variant  
 A:Reference number: A56795; MUID:93156788  
 A:Contents: Sprague-Dawley, kidneys  
 A:Accession: A56795  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-729 <KIN>  
 A:Cross-references: GB:SF4008; NID:9264804; PID:AA54274.1; PID:9264805  
 A:Note: sequence extracted from NCBI backbone (NCBIN:124259, NCBI:P:124270)  
 C:Superfamily: basic fibroblast growth factor receptor 1; Immunoglobulin homology; prote  
 C:Keywords: alternative splicing; ATP; growth factor receptor  
 F:80-14/Domain: immunoglobulin homology <IM>  
 F:383-668/Domain: protein kinase homology <KIN>  
 F:391-399/Region: protein kinase ATP-binding motif

alignment\_scores:  
 Quality: 364.00 Length: 246  
 Ratio: 2.167 Gaps: 6  
 Percent Similarity: 68.293 Percent Identity: 34.146

alignment\_block:  
 US-09-598-042a-2 x A56795 ..

Align seg 1/1 to: A56795 from: 1 to: 729

337 GACATTAGCCCGAGGAGAGAGACCTGGGGCCCGACACGTCCTCTGGGG 386  
 31 AspAlaLeuProSerSerGluAspAspAspAspAspAspSerSer 47  
 387 TCAGAG.....GACCCCGCAGCCAGCAGTGGGCGAGCCGCGCT 427  
 47 rGluGluLysGluThrAspAsnThrLysProAsnProValAlaProTyr 64  
 428 TCACACAGCCCTCAAGATGAGCGCGGGGTGATCGACGCGCGGTGGT 477  
 64 rPheSerProGluLysMetGluLysLysLeuHisAlaValProAlaAla 80  
 478 AGTCGCTGGGCTCAAGTGGCGGCGGCGGACCTCGGCCGACAT 527  
 81 LysThrValLysPheLysCysProSerSerGlyThrProSerProThrLe 97  
 528 CAGGTGATGAAGACACACAGGCTTGACGCGCCAGAG..... 567  
 97 uArgTrpLeuLysAsnGlyLysGluPhe...LysProAspHisArgIleG 113  
 568 ..GCCGCTGAGCCAGAGAGAGAGAGAGTGAAGTGAAGTGAAGAACTG 615

113 LysGlyTyrLysValArgTyrAlaThrTrpSerIleLeuMetAspSerVal 129  
 616 CCGCGCGAGAGACAGCGGCAATACACCTGCGGTGTCAGACCGCGGG 665  
 130 ValProSerAspLysGlyLysThrCysIleValGluAsnGluTyrGln 146  
 666 CGCATCAACAGCCACCTACAGAGTGGATGTGATCCAGCGAGCCGTTCCA 715  
 146 ySerIleAsnHisThrTyrGlnLeuAspValValGluArgSerProHisA 163  
 716 AGCCGCTGTCACAGGACAGCAGCCGCTGACACAGAGGTGACTTCGGG 765  
 163 rGProIleLeuGlnAlaGlyLeuProAlaAsnLysThrValAlaLeuGly 179  
 766 GGGACACAGTCCCTCCAGTGAAGGTGGCAGGAGACGTGAGCGGTGAT 815  
 180 SerAsnValGluPheMetCysLysValTyrSerAspProGlnProHisI 196  
 816 CCAGTGGCTGAAGCCGCTGAGTACGGCGCCGAGGCGCCACAACTCCA 865  
 196 eGlnTrpLeuLysHisIleGluVal.....AsnGlySerL 208  
 866 CCATCGATGTGGCGGCGCAGAACTTGTG...GTGCTGCCACAGGGTGAC 912  
 208 yslIleGlyProAspAsnLeuProTyrValGlnIleLeuLysThrAlaGly 224  
 913 GGTGTGTCGCGCGCGCCAGCGGTCTACATCAATCACTGCTCATACCGG 962  
 225 Val.....AsnThrThrAspLysGluMetGluValLeuHisIleArgAs 239  
 963 TGCCCGCAGAGAGATCGGCGATGTACATCTGCTTGGCGCCACACCA 1012  
 239 nValSerPheGluAspAlaGlyGluTyrThrCysLeuAlaGlyAsnSerL 256  
 1013 TGGGCTACAGCTTCGACGCGCTTCTTCAACGCTGTG 1050  
 256 leGlyLeuSerHisHisSerAlaTrpLeuThrValLeu 268



Tue Aug 21 14:45:53 2001

us-09-598-042a-2.rpt

---

Page 47

**THIS PAGE IS BLANK**



CC epithelial, stromal and blastocoele disorders, and inflammatory,  
CC angiogenic and immunologic disorders. AAC58242 to AAC58366 represent PCR  
CC primers and hybridisation probes used in the isolation of the human PRO  
CC sequences. AAC58367 to AAC58396 and AAB24057 to AAB24089 represent human  
CC PRO polynucleotide and protein sequences given in the exemplification of  
CC the present invention.

XX Sequence 504 AA:

alignment\_scores:

Quality: 2468.50 Length: 532  
Ratio: 5.208 Gaps: 3  
Percent Similarity: 89.098 Percent Identity: 88.722

alignment\_block:

us-09-598-042a-2 x AAB24066 ..

Align seg 1/1 to: AAB24066 from: 1 to: 504

```
28 CCGGCGCCGATGCGCGGATGCTTCATCCACGTGATGTCGCC 77
   |||||
21 ProAlaAlaAlaAlaArgGlyProProLysMetAlaAspLysValAlaPr 37
78 GGGTCGACGATTGCTGGCGAGTGGCGCGCTGGCGGCGACTATGGCGC 127
   |||||
37 o.....ArgLInValAlaArgLeuGlyArgThrValArgL 49
128 TGCAGTGGCCAGTGGAGGGGAGCGCGCGCGCTGACCATGTGACCAAG 177
   |||||
49 euGInCyProValGInGlyAspPropProLeuThrMetLTrpThrLys 65
178 GATGCGCCGACCATCCACAGCGCGTGGAGCGCTTCGCGTGGCGCGCA 227
   |||||
66 AspGlyArgThrLLeHisSerGlyTrpSerArgPheArgValLeuProG 82
228 GGGCTGAGGTGAAGCAGTGGAGCGGAGAGATGCCGCGTGAACGTGT 277
   |||||
82 nGlyLeuLysValLysGInValGInArgGInAspAlaGlyValLysValc 99
278 GCAAGGCCACCAAGCGCTTCGCGACCTGAGCGTCACTACACCTGCTC 327
   |||||
99 ySLyAlaLThrAsnGlyPheGlySerLeuSerValAsnLysTrpThrLeuVal 115
328 GTGCGGATGACATTAAGCCAGGAGGAGAGAGCTGGGCGCGGACAGCTC 377
   |||||
116 ValLeuAspAspLLeSerProGlyLysGInSerLeuGlyProAspSerSe 132
378 CTCTGGGGGTCAAGAGGAGCCCGCGACGACAGTGGGCGACGCGGCT 427
   |||||
132 rSerGlyGInGInLAspProAlaSerGInLTrpAlaArgProArgPr 149
428 TCACACAGCCCTCCAAAGATGAGGCGCGGGTGTATGACAGCGCGCTGT 477
   |||||
149 heThrGInProSerLysMetArgArgValLLeAlaArgProValGly 165
478 AGCTCGCGTGGCTCAAGTGGCTGGCGAGCGGACCGCTCGGCGCGACAT 527
   |||||
166 SerSerValArgLeuLysCysValAlaSerGlyHisProArgProAsp 182
528 CACGTGATGAAGAGCAGCAGCGCTTGACGCGCGGACAGCGCGCTGAGC 577
   |||||
182 eThrTrpMetLysAspAspGInAlaLeuThrArgProGInAlaAlaGIn 199
578 CCAGGAGAGAGAGTGAACACTGACGCTGAAGACCTGGCGCGGAGAGAC 627
   |||||
199 rOArgLysLysLysTrpThrLeuSerLeuLysAsnLeuArgProGInL 215
628 AGGCGCAATTAACCTGCGCGTGTGCAACCGCGCGGCGCGCATAGCGC 677
   |||||
216 SerGlyLysLysTrpThrCysArgValSerAsnArgAlaGlyAlaLLeAsnAl 232
678 CACCTAAGAGGTGATGTATCCAGCGGACCGCTTCAAGCCCGTGTCTCA 727
```

```
|||||
232 aThrTrpLysValAspValLLeGInArgThrArgSerLysProValLeuT 249
728 CAGGACGACGACCCCGTGAACACGACGCGTGGATTTGGGGGAGCACCGTCC 777
   |||||
249 hrcLysThrHisProValAsnThrThrValAspPheGlyLysThrHisThr 265
778 TTCACATGCAAGGTGGCGGACGACGAGTGAAGCGGTGATCACTGAGTGA 827
   |||||
266 PheGInCysLysValArgSerAspValLysProValLLeGInTrpLeuL 282
828 GCGCGTGAAGTACGGCGCGGAGGCGCGCACAACTCCACCATCATGTGCG 877
   |||||
282 sarGValGInLysGlyAlaGInGlyArgHisAsnSerThrLLeAspValG 299
878 GCGGCGCAAGATTGTGCTGCTGCCCGACGCGGTGACGTGCTGGCGGCGCC 927
   |||||
299 LysGInLysPheValValLeuProThrGlyAspValLTrpSerArgPro 315
928 GACGCTCCATCTCAATAAGTGTATCATCCCGCTGCGCGGACGACGAGA 977
   |||||
316 AspGlySerTrpLeuAsnLysLeuLeuLLeThrArgAlaArgGInAspAs 332
978 TCGGCGCATGTACATCTGCTTGGCGCGCAACACATGGGCTACAGTTCC 1027
   |||||
332 palagLysMetLysLLeCysLeuGlyAlaAsnThrMetGlyTrpSerPhea 349
1028 GCAAGCGCTTCCTCACCGTGTCCAGACCCCAAAACCGGACGAGCCACCT 1077
   |||||
349 rSerAlaPheLeuThrValLeuProAspProLysProProGlyProPro 365
1078 GTGGCTCTCTGTCTGTGGCGACTAGCTTGGCGCGGTGCTATCGG 1127
   |||||
366 ValAlaSerSerSerSerAlaThrSerLeuProTrpProValLLeG 382
1128 CATCCGACCGCGGCGTGTCTTCATCTGGGCGACCGCTGCTGGCTGT 1177
   |||||
382 yLLeProAlaGlyAlaValAlaPheLLeLeuGlyThrLeuLeuLTrpLeu 399
1178 GCCAGGCCAGAAAGACCGTGCACCCCGCGCTGCGCTGCCCTGCCCT 1227
   |||||
399 ySGInAlaGInLysLysProCysThrProAlaProAlaProProLeuPro 415
1228 GGGCAGCGCGCGCGCGGAGCGCGCGGACCGGACGCGGAGGAGACACCT 1277
   |||||
416 GLyHisArgProProGlyThrAlaArgAspArgSerGlyAspLysAspLe 432
1278 TTCCTGCTGGCGCGCTCAAGCGCTGCGCTGCTGCTGGGCTGTGTAGG 1327
   |||||
432 uProSerLeuAlaLLeuSerAlaGlyProGlyValGlyLeu..... 446
1328 AGCATGGGCTCCGCGGACCGCCCGACGACCTACTGGGCGGACCGCAGTT 1377
   |||||
446 ..... 446
1378 GCTGGCCCTAAGTTGTATCCCAAACTTACACAGACATCCACACACAC 1427
   |||||
446 ..... 446
1428 ACACTGTATTGCGCGCGCTGTGTGAGAGACATGGGCTCCGCGAGCGCC 1477
   |||||
447 .....Cys..GInGInHisGlySerProAlaAlaLys 456
1478 CCAGCACTTACTGGGCGGAGCGCGGATGCTGGCGCTTAAGTTGTACCCCA 1527
   |||||
456 oGInHisLeuLeuGlyProGlyProValAlaGlyProLysLeuLysProL 473
1528 AACTCTACACAGACATCCACACACACACACACACACACACACACACAC 1577
   |||||
473 ySLeLysTrpAspLLeHisThrHisThrHisThrHisSerHisThrHis 489
1578 TCACACGTGAGGAGGCAAGTGCACACGACATCCACTATCATGCTGC 1622
   |||||
```

490 SerHisValGluGlyValHisGlnHisIleHisTyrGlnCys 504

seq.name: /SIS1/gcdata/geneseq/geneseq/AA2000.DAT:AAV66656

seq\_documentation\_block:

ID AAV66656 standard; protein: 504 AA.

XX AAV66656;

AC

XX 05-APR-2000 (first entry)

DT

XX Membrane-bound protein PRO943.

DE

XX Membrane-bound polypeptide: PRO polypeptide; LDL receptor; TIE ligand;

KM pharmaceutical; receptor Immunoadhesin; gene mapping.

XX

OS Homo sapiens.

XX

PN WO963088-A2.

XX

PD 09-DEC-1999.

XX

PF 02-JUN-1999; 99WO-US12252.

XX

PR 02-JUN-1998; 98US-0087607.

PR

PR 02-JUN-1998; 98US-0087609.

PR

PR 02-JUN-1998; 98US-0087759.

PR

PR 03-JUN-1998; 98US-0087827.

PR

PR 04-JUN-1998; 98US-0088021.

PR

PR 04-JUN-1998; 98US-0088025.

PR

PR 04-JUN-1998; 98US-0088028.

PR

PR 04-JUN-1998; 98US-0088029.

PR

PR 04-JUN-1998; 98US-0088030.

PR

PR 04-JUN-1998; 98US-0088033.

PR

PR 04-JUN-1998; 98US-0088036.

PR

PR 05-JUN-1998; 98US-0088167.

PR

PR 05-JUN-1998; 98US-0088202.

PR

PR 05-JUN-1998; 98US-0088217.

PR

PR 05-JUN-1998; 98US-0088655.

PR

PR 10-JUN-1998; 98US-0088722.

PR

PR 10-JUN-1998; 98US-0088730.

PR

PR 10-JUN-1998; 98US-0088734.

PR

PR 10-JUN-1998; 98US-0088738.

PR

PR 10-JUN-1998; 98US-0088740.

PR

PR 10-JUN-1998; 98US-0088741.

PR

PR 10-JUN-1998; 98US-0088742.

PR

PR 10-JUN-1998; 98US-0088810.

PR

PR 10-JUN-1998; 98US-0088811.

PR

PR 10-JUN-1998; 98US-0088824.

PR

PR 10-JUN-1998; 98US-0088825.

PR

PR 10-JUN-1998; 98US-0088826.

PR

PR 11-JUN-1998; 98US-0088858.

PR

PR 11-JUN-1998; 98US-0088861.

PR

PR 11-JUN-1998; 98US-0088863.

PR

PR 11-JUN-1998; 98US-0088876.

PR

PR 12-JUN-1998; 98US-0089090.

PR

PR 12-JUN-1998; 98US-0089105.

PR

PR 16-JUN-1998; 98US-0089440.

PR

PR 16-JUN-1998; 98US-0089512.

PR

PR 16-JUN-1998; 98US-0089514.

PR

PR 17-JUN-1998; 98US-0089532.

PR

PR 17-JUN-1998; 98US-0089538.

PR

PR 17-JUN-1998; 98US-0089558.

PR

PR 17-JUN-1998; 98US-0089599.

PR

PR 17-JUN-1998; 98US-0089600.

PR

PR 17-JUN-1998; 98US-0089653.

PR

PR 18-JUN-1998; 98US-0089801.

PR

PR 18-JUN-1998; 98US-0089807.

PR

PR 18-JUN-1998; 98US-0089901.

PR

PR 18-JUN-1998; 98US-0089908.

PR

PR 19-JUN-1998; 98US-0089947.

PR

PR 19-JUN-1998; 98US-0089948.

PR

PR 19-JUN-1998; 98US-0089952.

PR

PR 22-JUN-1998; 98US-0090246.  
PR 22-JUN-1998; 98US-0090252.  
PR 22-JUN-1998; 98US-0090254.  
PR 23-JUN-1998; 98US-0090349.  
PR 23-JUN-1998; 98US-0090355.  
PR 24-JUN-1998; 98US-0090429.  
PR 24-JUN-1998; 98US-0090431.  
PR 24-JUN-1998; 98US-0090435.  
PR 24-JUN-1998; 98US-0090444.  
PR 24-JUN-1998; 98US-0090445.  
PR 24-JUN-1998; 98US-0090461.  
PR 24-JUN-1998; 98US-0090472.  
PR 24-JUN-1998; 98US-0090535.  
PR 24-JUN-1998; 98US-0090538.  
PR 24-JUN-1998; 98US-0090540.  
PR 24-JUN-1998; 98US-0090557.  
PR 25-JUN-1998; 98US-0090676.  
PR 25-JUN-1998; 98US-0090688.  
PR 25-JUN-1998; 98US-0090680.  
PR 25-JUN-1998; 98US-0090691.  
PR 25-JUN-1998; 98US-0090694.  
PR 25-JUN-1998; 98US-0090695.  
PR 25-JUN-1998; 98US-0090696.  
PR 26-JUN-1998; 98US-0090862.  
PR 26-JUN-1998; 98US-0090863.  
PR 01-JUL-1998; 98US-0091358.  
PR 01-JUL-1998; 98US-0091360.  
PR 02-JUL-1998; 98US-0091544.  
PR 02-JUL-1998; 98US-0091478.  
PR 02-JUL-1998; 98US-0091486.  
PR 02-JUL-1998; 98US-0091519.  
PR 02-JUL-1998; 98US-0091626.  
PR 02-JUL-1998; 98US-0091628.  
PR 02-JUL-1998; 98US-0091633.  
PR 02-JUL-1998; 98US-0091646.  
PR 02-JUL-1998; 98US-0091673.  
PR 07-JUL-1998; 98US-0091978.  
PR 07-JUL-1998; 98US-0091982.  
PR 09-JUL-1998; 98US-0092162.  
PR 10-JUL-1998; 98US-0092472.  
PR 20-JUL-1998; 98US-0093339.  
PR 30-JUL-1998; 98US-0094651.  
PR 04-AUG-1998; 98US-0095282.  
PR 04-AUG-1998; 98US-0095285.  
PR 04-AUG-1998; 98US-0095301.  
PR 04-AUG-1998; 98US-0095302.  
PR 04-AUG-1998; 98US-0095318.  
PR 04-AUG-1998; 98US-0095321.  
PR 04-AUG-1998; 98US-0095325.  
PR 10-AUG-1998; 98US-0095915.  
PR 10-AUG-1998; 98US-0095929.  
PR 10-AUG-1998; 98US-0096012.  
PR 11-AUG-1998; 98US-0096143.  
PR 11-AUG-1998; 98US-0096146.  
PR 12-AUG-1998; 98US-0096329.  
PR 17-AUG-1998; 98US-0096757.  
PR 17-AUG-1998; 98US-0096766.  
PR 17-AUG-1998; 98US-0096768.  
PR 17-AUG-1998; 98US-0096773.  
PR 17-AUG-1998; 98US-0096791.  
PR 17-AUG-1998; 98US-0096867.  
PR 17-AUG-1998; 98US-0096891.  
PR 17-AUG-1998; 98US-0096894.  
PR 17-AUG-1998; 98US-0096895.  
PR 17-AUG-1998; 98US-0096897.  
PR 18-AUG-1998; 98US-0096949.  
PR 18-AUG-1998; 98US-0096950.  
PR 18-AUG-1998; 98US-0096959.  
PR 18-AUG-1998; 98US-0096960.  
PR 18-AUG-1998; 98US-0097022.  
PR 19-AUG-1998; 98US-0097141.  
PR 20-AUG-1998; 98US-0097218.



```

1178 GCCAGGCCAGAGAAGCCGTGACACCCCGCGCTGCCCTCCCTGCT 1227
      |||||
399 y6GlnAlaInLysProCysThrProAlaProAlaProLeuPro 415
1228 GGGACACCGCCCGCGGGAGCGCCCGCAGCCGAGCGAGACAGACCT 1277
      |||||
416 GlyHisArgProProGlyThrAlaArgAspArgSerGlyAspLysAsp 432
1278 TCCCTGCTGGCCCGCCCTGACGCGCTGGCCCTGCTGGCGCTGTGAGG 1327
      |||||
432 uProSerLeuAlaAlaLeuSerAlaGlyProGlyAlaGlyLeu..... 446
1328 AACATGGGTCTCGCGAGCCCGCCAGCACTACTGGGCGCCAGCCAGTT 1377
446 ..... 446
1378 GCTGGCCCTTAAGTGTATGCCCAAACTCTACAGACATCCACACACAC 1427
446 ..... 446
1428 AACTGTATTCGCGCCGCTGTGTGAGAGCATGGGTCTCCGCGACGCC 1477
      ||| |||||
447 .....Cys..GluGlnHisGlySerProAlaAlaPr 456
1478 CCAGCACTACTGGGCCGCGCCAGCCAGTTGCTGGCCCTAAGTTGTACCC 1527
      |||||
456 oGlnHisLeuLeuGlyProGlyProValAlaGlyProLysLeuTyProL 473
1528 AACTCTACAGACATCCACACACACACACACATCTCTACACACAC 1577
      |||||
473 yLeuTyThrAspIleHisThrHisThrHisThrHisThrHis 489
1578 TCACACGTGGAGGCGCAAGGTCCACACGACATCCATCATCAGTGC 1622
      |||||
490 SerHisValGluGlyLysValHisGlnHisIleHisTyGlnCys 504

seq_name: /SISD1/gcdata/geneseq/genesep/AA2001.DAT: AAB65179
seq_documentation_block:
ID AAB65179 standard; Protein: 504 AA.
XX
AC AAB65179;
XX
DT 02-APR-2001 (first entry)
XX
DE Human PRO943 (UNQ480) protein sequence SEQ ID NO:119.
XX
KW Human; secreted and transmembrane protein; PRO; cytosolic;
KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;
KW diagnostic assay.
XX
OS Homo sapiens.
XX
PN WO200073454-A1.
XX
PD 07-DEC-2000.
XX
PE 30-MAR-2000; 2000WO-US08439.
XX
PR 02-JUN-1999; 99WO-US12252.
PR 23-JUN-1999; 99US-0141037.
PR 07-JUL-1999; 99US-0143048.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0146222.
PR 17-AUG-1999; 99US-0149386.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 08-OCT-1999; 99US-0158663.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 16-DEC-1999; 99WO-US30095.

```

```

PR 20-DEC-1999; 99WO-US30911.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06864.
PR 20-MAR-2000; 2000WO-US07377.
XX
PA (GETH ) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI Ferrara N, Fong S, Gerber H, Gertlisen ME, Goddard A, Godowski PJ;
PI Grimaldi CJ, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF;
PI Roy MA, Stewart TA, Tumes D, Watanabe CK, Williams PM, Wood WI;
PI Zhang Z;
XX
DR WPI: 2001-032160/04.
DR N-PSDB: AAF44130.
XX
PT PRO polynucleotides used to produce polypeptides used to target
PT bioactive molecules such as toxins, radiolabels or antibodies, to
PT specific cells, to cause targeted cell death.
XX
PS Claim 12; Fig 70; 935pp: English.
XX

```

The present invention describes human secreted and transmembrane PRO proteins. The PRO proteins have cytostatic activity. The PRO proteins can be used for targeted delivery of bioactive molecules, such as toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide sequences, and their fragments, can be used as hybridisation probes, in chromosomal and gene mapping, and in the generation of anti-sense RNA and DNA. They may also be used to produce transgenic animals which are used to develop and screen therapeutically useful reagents. The PRO nucleotide and protein sequence can be used for tissue typing and in treating cancer. Anti-PRO antibodies can be used in diagnostic assays. CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used in the isolation of human PRO sequences. AAF44087 to AAF44269 and CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein sequences given in the exemplification of the present invention.

Sequence 504 AA;

alignment\_scores:  
 Quality: 2468.50 Length: 532  
 Ratio: 5.208 Gaps: 3  
 Percent Similarity: 89.098 Percent Identity: 88.722

alignment\_block:  
 US-09-598-042a-2 x AAB65179 ..

Align seg 1/1 to: AAB65179 from: 1 to: 504

```

28 CCGGCGCCATGCGCGCGGATTTGCTTCATCCACGATGATGCTTCC 77
      ||||| ||| |||
21 ProAlaAlaAlaAlaArgGlyProProLysMetAlaAspLysValValPr 37
78 GGGTCGACGATTTCTGCGCAGGTGGCGCGGCTGGCGCGCACATGCGGC 127
      |||||
37 o.....ArgGlnValAlaArgLeuGlyArgThrValArgL 49
128 TGCAGTGGCCAGTGGAGGGGAGACCCCGCCGCTGACCATGTGGACAG 177
      |||||
49 euGlnCysProValGluGlyAspProProLeuThrMetTrpTrpLys 65
178 GATGGCGGACCATCCACAGCGGCTGAGCCGCTTCGCGGCTGCGCGCA 227
      |||||
66 AspGlyArgThrIleHisSerGlyTrpSerArgPheArgValLeuProL 82

```

```

228 GGGCTGAGGTGAGAGGTGAGAGGAGATGCCGGCTGTACGTGT 277
      |||
82  nGlyLeuLysValLysGlnValGlnArgGlnAspAlaGlyValIleTyrAlaC 99
278 GCAAGGCCAACAGCGCTTGGCAGCCTGAGCGTCAACTACACCTCGTC 327
      |||
99  ySLysAlaThrAsnGlyPheGlySerLeuSerValAsnTyrThrLeuVal 115
328 GTGCTGATGATTAAGCCCGAGAGAGAGCGCTGGGGCCCGCAGAGTC 377
      |||
116 ValLeuAspPheLeuSerProGlyLysGlnSerLeuGlyProAspSerSe 132
378 CTCTGGGGGTCAAGAGAGACCCCGCAGCAGACAGTGGGACAGCGCGCT 427
      |||
132 rSerGlyGlyGlnGlnAspProAlaSerGlnGlnThrAlaArgProArgP 149
428 TCACAGACCCCTCCAGATGAGGGCGGGGTGATCGCACGGCCGTGGGT 477
      |||
149 heThrGlnProSerLysMetArgArgValIleAlaIleArgProValGly 165
478 AGCTCCGTGGCGTCAAGTGGGTGGCCAGCGGGGACCCCTGGCCGACAT 527
      |||
166 SerSerValAlaGlyLeuLysCysValAlaSerGlyHisProArgProAsp 182
528 CACGTGATGAGAGAGACAGCAGCGCTTGACCGCCAGAGGCGCGCTGAGC 577
      |||
182 eThrTrpMetLysAspAspGlnAlaLeuThrArgProGlnAlaIleGln 199
578 CCAGGAGAGAGAACTGGACACTGAGCCTGAAGAACCTGCGCGCGGAGAG 627
      |||
199 roArgLysLysLysTrpThrLeuSerLeuLysAsnLeuArgProGlnAsp 215
628 AGCGGCAAAATACACCTGCGCGCTGCGAACCGCGGGCGGCATCAACGC 677
      |||
216 SerLysLysTrpThrCysArgValSerAsnArgAlaGlyAlaIleAsnAl 232
678 CACCTACAGGTGATGTATGCACGCGGACCGCTTCCAAAGCCGTGTCA 727
      |||
232 aThrTyrLysValAspValIleGlnArgThrArgSerLysProValLeu 249
728 CAGGACGAGCACCGCGTGACAGCAGCGGTGACTCGGGGGACACACGTCC 777
      |||
249 hrGlyThrHisProValAsnThrThrValAspPheGlyTyrThrThrSer 265
778 TTCCAGTGCAGAGTGCAGCAGCGAGCTGAAGCGGTGATCCAGTGGTGA 827
      |||
266 PheGlnCysLysValArgSerAspValLysProValIleGlnTrpLeuL 282
828 GCGCGTGAGTACGCGCGCGAGCGCGCACACTCCACATCGATGTGG 877
      |||
282 sarValaGlnTyrGlyAlaGlnGlyArgHisAsnSerThrIleAspValG 299
878 GCGCGCAGAGATTGTGTGCTGCCAGCGGTGACGTGTGCTGCCGGCCC 927
      |||
299 LysGlnLysLysPheValValLeuProThrGlyAspValTyrPheArgPro 315
928 GACGCTCTACTCAATAAAGCTGCTCATCACCGCTGCCCGCCGACAGAGA 977
      |||
316 AspGlySerTyrLeuAsnLysLeuLeuIleThrArgAlaArgGlnAspAs 332
978 TGGCGGATGATGATGCTTGGCGCGCACACACATGGGCTACAGCTTCC 1027
      |||
332 palaglyMetLysTyrIleCysLeuGlyAlaAsnThrMetGlyTyrSerPhe 349
1028 GCAGCGCTTCTCAGCGGTGCTGCCAGAACCAACGCGCAGGCGCACCT 1077
      |||
349 rGSerAlaPheLeuThrValLeuProAspProLysProProGlyProPro 365
1078 GTGGCTCTCTCTCTCTGCGCACAGCTGCGCGCGCGCGGTGATCGG 1127
      |||
366 ValAlaSerSerSerSerAlaThrSerLeuProThrProValValIleG 382
1128 CATCCAGCGCGCGCTGTTCATCTGAGGACCCCTGCTGTGCTTT 1177

```

```

      |||
382 yLleProAlaGlyAlaValPheIleLeuGlyThrLeuLeuThrPheL 399
1178 GCCAGGCCAGAGAGAGCGGTGACACCCCGCGCTGCCCTCCCTGCT 1227
      |||
399 ySglnAlaGlnLysLysProCysThrProAlaProAlaProLeuPro 415
1228 GGGCAGCGCGCGCGGGGAGCGCGGAGCGGACCGCAGCGAGACAGGACT 1277
      |||
416 GlnHisArgProProGlyThrAlaArgAspArgSerGlyAspLysAspLe 432
1278 TCCCTGCTGGCGCGCGCTCAGCGCTGGCGCGCTGCTGTGGGCTGTGAGG 1327
      |||
432 uProSerLeuAlaIleLeuSerAlaGlyProGlyValGlyLeu..... 446
1328 AGCATGGGTCTCCGGCAGACCCCGCAGCATTACTGGGCGCCAGGCCATT 1377
      |||
446 ..... 446
1378 GCTGGCCCTAAGTTGTACCCCAACTCTACAGACATCCACACACAC 1427
      |||
446 ..... 446
1428 ACACTGATTGCGGCGCGCTGTGTGAGAGACATGGTCTCCGCGAGCCC 1477
      |||
447 .....Cys..GlnGlnHisGlySerProAlaIleArg 456
1478 CCAGCACTTACTGGCGCGCGCGCGCGCGCTGCTGGCGCGCTGATGTA 1527
      |||
456 oglnHisLeuLeuGlyProGlyProValAlaGlyProLysLeuTyrPro 473
1528 AACTCTACACAGACATCCACACACACACACACACACTCCACACACAC 1577
      |||
473 ySLeuTyrThrAspIleHisThrHisThrHisThrHisSerHisThrHis 489
1578 TCACAGGTGAGGCGCAAGTCCACAGCAGACATCCAGTATCATGTC 1622
      |||
490 SerHisValGlnGlyLysValHisGlnHisIleHisTyrGlnCys 504
seq_name: /stds1/gcgdata/geneseq/geneseq/AA2001.DAT:AA66264
seq_documentation_block:
ID   AAB66264 standard; Protein; 504 AA.
XX
AC   AAB66264;
XX
DT   05-APR-2001 (first entry)
XX
DE   Human MANGO 003 SEQ ID NO: 5.
XX
KW   Membrane associated protein; secreted protein; human; mouse; rat;
KW   INTERCEPT 340; MANGO 003; MANGO 347; TANGO 272; TANGO 354;
KW   TANGO 378; skeletal disorder; cardiovascular disorder; renal disorder;
KW   haematopoietic disorder; neural disorder; hepatic disorder;
KW   neoplastic disease.
XX
OS   Homo sapiens.
XX
PN   WO200100673-A1.
XX
PD   04-JAN-2001.
XX
PF   29-JUN-2000; 2000MO-US18198.
XX
PR   30-JUN-1999; 99US-0345464.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Barnes TM, Fraser CC, Wighton N, Myers P, Busfield SJ, Sharp JD;
DR   WPT: 2001-050128/06.
XX
N-PSDB: AAF27781.
XX

```



PT Isolated secreted or transmembrane proteins are used for diagnosis and  
PT treatment of neoplastic and haematopoietic disorders e.g. T cell  
PT disorders, cancer and tumours -

XX Claim 9: Page 216-217; 294pp; English.

XX The present invention provides the protein and coding sequences for a  
CC number of membrane associated and secreted proteins from human, mouse and  
CC rat. The proteins are designated INTERCEPT 340, MANGO 003, MANGO 347,  
CC TANGO 272, TANGO 295, TANGO 254 and TANGO 378. The proteins are all  
CC involved in signal transduction and the sequences can be used in the  
CC treatment of cardiovascular, renal, hepatic, neural, neoplastic, skeletal  
CC and haematopoietic disorders.

XX Sequence 504 AA:

Alignment\_scores:

Quality: 2468.50 Length: 512  
Ratio: 5.208 Gaps: 3  
Percent Similarity: 89.098 Percent Identity: 88.722

Alignment\_block:

US-09-598-042A-2 x AAB66264 ..

Align seg 1/1 to: AAB66264 from: 1 to: 504

```
28 CCGCGCGCCATGCGCGGATTCCTTCATCCAGTGATGTCGTCCTCC 77
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
21 ProAlaAlaAlaAlaArgGlyProProLysMetAlaAspLysValAlaPr 37
78 GGGTCGACAGATTTCGTCGACAGTGGCCGCTGGCCGCTACATAGCGC 127
|  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
37 o.....ArgGlnValAlaArgLeuGlyArgThrValArgL 49
128 TGCAGTGCCTGAGGGGAGCGCGCGCGCTGACATGTGAGCAAG 177
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
49 euGlnCysProValIgluGlnAspProProProLeuThrMetIrrhLys 65
178 GATGGCCGACCATCCACAGCGGTGAGACCGCTCCGCGTCTGCGCGA 227
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
66 AspGlyArgThrIleHisSerGlyTrpSerArgPheArgValLeuProI 82
228 GGGCGGTGAAGTGAAGCAGGTGAGCGGGAGAGATGCGCGTGTACGT 277
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
82 nGlyLeuLysValLysGlnValIgluArgLysPalaGlyValIlyValC 99
278 GCAAGCCCAACCAAGCGCTTCGAGCGCTGAGCTCAACTACACCGTC 327
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
99 yslYAlaIThrAsnGlyPheGlySerLeuSerValAsnIlyrThrLeuVal 115
328 GTGCTGATGACATTAGCCAGGAGAGAGAGACCTGGGCGCCGACAGCT 377
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
116 ValLeuAspAspIleSerProGlyLysGlnSerLeuGlyProAspSerSe 132
378 CTCTGGGGGTCAAGAGACCGCGCCAGCAGCAGTGGGAGACCGCGCT 427
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
132 rSerIlyGlnGlnAspProAlaSerGlnGlnIrrPalaArgProAlaPr 149
428 TCACACAGCCCTCCAAAGATGAGCGCGCGGTATGACAGCGCGCGGT 477
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
149 heThrGlnProSerLysMetArgArgValIleAlaArgProValIglY 165
478 AGCTCGGTGGGCTCAAGTGCCTGGCCAGCGGGCAACCTCGGCGCGACAT 527
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
166 SerSerValArgLeuLysCysValAlaSerGlyHisProArgProAspI 182
528 CACGTGATGAAGAGACGACCGCCTTGACGCGCCAGAGCGCGGTGAGC 577
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
182 eThrTrpMetLysAspAspGlnAlaLeuThrArgProGlnAlaAlaGlnP 199
578 CCAGGAAGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 627
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

```
199 roArgLysLysLysTrpThrLeuSerLeuLysAsnLeuArgProGlnAsp 215
628 ACCGGCAATFACCTGCGCGCTGTGCAACCGCGGGCGCATACGCC 677
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
216 SerGlyLysTrpThrCysArgValSerAsnArgAlaGlyAlaIleAsnAl 232
678 CACGTACAGGAGTGTGATCCAGGGAGCGCGCTCCCAAGCGCGGTCTA 727
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
232 aThrTrpLysValAspValIleGlnIrrThrArgSerLysProValLeuT 249
728 CAGGACAGCAGCCCGTGAACAGCAGCGTGAAGTGGGGGAGACCAAGTCC 777
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
249 hGclYThrHisProValAsnThrThrValAspPheGlyLyrThrHisSer 265
778 TTCAGTGCAGGTGGCGGACGACGAGTGAACCGGTGATCCAGTGCCTGA 827
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
266 PheGlnCysLysValArgSerAspValLysProValIleGlnIrrLeuL 282
828 GCGCGGTGAGTACGGCGCGCGAGCGCGCCCAACCTCCACATCGATGTGG 877
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
282 sarGValGlnTrpGlyAlaGlnIlyArgHisAsnSerThrIleAspValG 299
878 GCGCCAGAAAGTTTGGTGTGTGCTGCCACGCGGTGACGTGTGTGCGGCC 927
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
299 lylGlnLysPheValValLeuProThrGlyAspValIrrSerArgPro 315
928 GACGCTCTTACCTCAATAGCTGCTCATACCGCTGCGCGCGAGAGCA 977
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
316 AspGlySerTrpLeuAsnLysLeuLeuIleThrArgAlaArgLysAsps 332
978 TCGCGGCATGATCATCTGCTTGGCGCCCAACACATGGGCTACAGTTC 1027
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
332 palAglMetTrpIleCysLeuGlyAlaAsnThrMetGlyTrpSerPhe 349
1028 GCAGCGCTTCTTCAACCGTGTGCGCAGACCAAAACCGCGAGGCACT 1077
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
349 rSerAlaPheLeuThrValLeuProAspProLysProProGlyProPro 365
1078 GTGGCTCTCTGCTCTGCGCACATAGCTGCGCGTGGCGCGGTGCATCGG 1127
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
366 ValAlaSerSerSerSerIleThrSerLeuProTrpProValIleGln 382
1128 CATCCAGCGCGCGCTGCTTCATCTCGGGCAACCGCTGCTGCTGCTT 1177
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
382 ylleProAlaGlyAlaValPheIleLeuGlyThrLeuLeuLeuTrpLeuC 399
1178 GCCAGGCGCAGAAAGCGGTGACACCGCGCGCTGCGCGCTGCGCT 1227
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
399 yslGlnAlaGlnLysLysProCysThrProAlaProAlaProProLeuPro 415
1228 GGGCACCGCGCGCGGGAGCGCGCGCGACCGCAGCGGAGACAGACCT 1277
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
416 GlyHisArgProProGlyThrAlaArgAspArgSerGlyAspLysAspLe 432
1278 TCCGTCGTGGCGCGCGCTTCAGCGCGCGCTGCTGTTGGGGCTGTGAGG 1327
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
432 uProSerLeuAlaAlaLeuSerIleGlyProGlyValGlyLeu..... 446
1328 AGCATAGGTCTCCGGCAGCGCGCGCACTTACTGGGCGCAGCGCACTT 1377
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
446 ..... 446
1378 GCTGGCCCTAAGTTGATCCCAAACCTTACACAGACATCCACACACAC 1427
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
446 ..... 446
1428 ACACTGTATTGGCGCGCGCTGTGTGAGGAGCATGGGTCTCGGCGAGCCC 1477
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
447 .....Cys..GluGlnHisGlySerProAlaIleAlaPr 456
1478 CCAGCACTTACTGGCGCGCAGCGCAAGTGTGCGCGCTAAGTTGTAACCCA 1527
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
456 oGlnHisLeuLeuGlyProGlyProValAlaGlyProLysLeuTrpProL 473
```

```
1528 AACCTACACAGACATCCACACACACACACACTCTCACACAC 1577
|||||
473 YSLuTYrThrAspIleHisThrHisThrHisSerHisThrHis 489
1578 TCACACGCGAGGCGACGTCACACACATCCAGTCGTC 1622
|||||
490 SerHisValGluGlyValHisGlnHisHisThrHisGlyGly 504

seq_name: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT.AAY92864

seq_documentation_block:
ID AAY92864 standard; protein; 504 AA.
AC AAY92864;
XX
XX 29-AUG-2000 (first entry)
XX
DE Human fibroblast growth factor receptor 5.
XX
XX FGFR-5; fibroblast growth factor receptor 5; cyrostatic; anti-sclerotic;
XX immunomodulatory; gastrointestinal; virucide; anti-inflammatory;
XX anti-ischemic; anti-atherosclerosis; angiogenic; endocrine;
XX anti-diabetic; gene therapy.
XX
OS Homo sapiens.
XX
XX Key
XX Location/Qualifiers
XX 1..24
XX Peptide
XX /label= leader_sequence
XX 23..37
XX Peptide
XX /label= antigenic
XX 25..504
XX Protein
XX /label= mature-protein
XX 25..117
XX Domain
XX /label= extracellular
XX /note= "immunoglobulin domain I"
XX 39..48
XX Peptide
XX /label= antigenic
XX 51..59
XX Peptide
XX /label= antigenic
XX 62..76
XX Peptide
XX /label= antigenic
XX 81..97
XX Peptide
XX /label= antigenic
XX 101..104
XX Peptide
XX /label= antigenic
XX 118..143
XX Domain
XX /label= acid_box_domain
XX 119..170
XX Peptide
XX /label= antigenic
XX 144..239
XX Domain
XX /label= extracellular
XX /note= "immunoglobulin domain II"
XX 176..204
XX Peptide
XX /label= antigenic
XX 209..228
XX Peptide
XX /label= antigenic
XX 238..247
XX Peptide
XX /label= antigenic
XX 240..357
XX Domain
XX /label= antigenic
XX /label= extracellular
XX /note= "immunoglobulin domain III"
XX 259..262
XX Peptide
XX /label= antigenic
XX 268..275
XX Peptide
XX /label= antigenic
XX 282..302
XX Peptide
XX /label= antigenic
XX 307..320
XX Peptide
XX /label= antigenic
XX 326..334
XX Peptide
XX /label= antigenic
```

```
FT Peptide 356..375
FT /label= antigenic
FT Domain 358..373
FT /label= membrane_proximal_domain
FT Domain 374..403
FT /label= transmembrane_domain
FT Peptide 401..435
FT /label= antigenic
FT Domain 404..504
FT /label= intracellular_domain
FT Peptide 440..443
FT /label= antigenic
FT Peptide 446..455
FT /label= antigenic
FT Peptide 462..475
FT /label= antigenic
FT Peptide 483..496
FT /label= antigenic
XX
XX W0200024756-A1.
XX
XX PD 04-MAY-2000.
XX
XX PF 17-JUN-1999; 99WO-US13620.
XX
XX PR 23-OCT-1998; 98US-0105465.
XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Ruben SM, Young PE;
XX
XX DR WPI: 2000-387035/33.
XX N-PSDB; AAA28842.
XX
XX PT Nucleic acids encoding fibroblast growth factor-5 useful for the
XX prevention, diagnosis and treatment of conditions associated with
XX tissue repair and aberrant cell functions, e.g. cell survival and
XX proliferation
XX
XX PS Claim 11; Flg 1A-C; 182pp; English.
XX
XX CC This is the fibroblast growth factor receptor protein, FGFR-5. The
XX FGFR-5 protein and DNA may be used in the prevention, treatment and
XX diagnosis of diseases or conditions associated with inappropriate FGFR-5
XX expression and activity. For example, the nucleic acids (and vectors
XX containing them) and the FGFR-5 polypeptide may be used to treat
XX disorders associated with increased or decreased cell survival (such as
XX cancers (e.g. leukemia, colonic cancer, testicular cancer and follicular
XX lymphomas), autoimmune disorders (e.g. multiple sclerosis and Crohn's
XX disease) viral infections (e.g. herpes viruses), inflammation, graft
XX versus host disease, acute and chronic graft rejection, ischemic injuries
XX and atherosclerosis), activation, secretion, migration, differentiation
XX and proliferation, diseases associated with defects in wound healing,
XX mucositis, defects of angiogenesis, immune dysfunction, endocrine
XX dysfunction and insulin secretion disorders. Anti-FGFR-5 antibodies may
XX also be used as diagnostic agents for detecting the presence of FGFR-5
XX polypeptides in samples.
XX
XX SO Sequence 504 AA:

alignment_scores:
Quality: 2453.50 Length: 532
Ratio: 5.198 Gaps: 3
Percent Similarity: 88.722 Percent Identity: 88.346

alignment_block:
us-09-598-042a-2 x AAY92864 ..
Align seg 1/1 to: AAY92864 from: 1 to: 504
28 CCGCGCGCATGGCGGCGGATGTCCTCATCCAGCATGATGTCCTCC 77
||||| ||| |||
```

```

21 ProAlaAlaAlaAlaArgGlyProProLysMetAlaAspLysValValPr 37
78 GGGTCACAGATTTCGTGGCAGGTGGCCGGCTGGCCGACATATGGCGC 127
37 O.....ArgGlnValAlaAlaArgLeuGlyArgThrValArgL 49
128 TGCAGTGGCCCGCATGGAGGGGGAGCCCGCGCGGTGACCATGTGGACACAG 177
49 euGlnCysProValGluGlnLysProProProLeuThrMetTrpThrLys 65
178 GATGGCCGCACCATCATCAGCGCGCTGAGACCGCTTCCGCTGTGCGCCA 227
66 AspGlyArgThrIleHisSerGlyTrpSerArgThrValValLeuProG 82
228 GGGCTGTAAGTGAAGCAGGTGAGCGGAGGAGATGCGCGCTGTACGTGT 277
82 nGlyLeuLysValLysGlnValGluArgGluAspAlaGlyValTyrValC 99
278 GCAGGGCCACCAAGGGCTTGGGAGCGCTGAGCGTCATACACCGCTGTC 327
99 yslYsAlaThrAsnGlyPheGlySerLeuSerValAsnTyrTrpLeuVal 115
328 GTGCTGATGACATTAGCCACGAGGAGAGAGAGCTGGGGCCGACAGCTC 377
116 ValLeuAspAspIleSerProGlyLysGluSerLeuGlyProAspSerSe 132
378 CTCGGGGGGTAAAGAGACCCCGCCAGCCAGCTAGTGGGACGACCGGCT 427
132 rSerGlyGlyGlnGlnLysProAlaSerGlnGlnTrpAlaArgProArgP 149
428 TCACACAGCCCTCCAGATGAGGGCGGGGTGATGCGGCGCGCGCTGGGT 477
149 heThrGlnProSerLysMetArgArgValIleAlaArgProValGly 165
478 AGCTCCGTGCGCTCAAGTGGCTGGCCAGCGGAGACCTCGCGCCGACAT 527
166 SerSerValArgLeuLysCysValAlaSerGlyHisProArgProAspI 182
528 CAGGTGATGAAGACACAGCGCTTGACCGCCCGACAGGCGCGTGAAC 577
182 eThrTrpMetLysAspArgGlnAlaLeuThrArgProGluAlaAlaGluP 199
578 CCAGAGAGAGAGAGTGGGACTGAGCCTGGAAGAACCTGGCGCGGAGGAC 627
199 roArgLysLysLysTrpThrLeuSerLeuLysAsnLeuArgProGluAsp 215
628 AGCGCAAAATACACCTGCGCGCTGCAAAACCGCGCGCGCCATCAAGCG 677
216 SerGlyLysTyrThrCysArgValSerAsnArgAlaGlyAlaIleAsnAl 232
678 CACCTACAAAGTGTGATGCCAGCGGACCGGTCCAAAGCCGTGCTCA 727
232 aThrTyrLysValAspValIleGlnArgThrArgSerLysProValLeuT 249
728 CAGGACGCGCACCCCGTGAACAGAGCGGTGACTCGGGGGAGACACAGTCC 777
249 hrGlyThrHisProValAsnThrThrValAspPheGlyGlyThrThrSer 265
778 TTCCAGTGCAGAGTGCAGCAGCAGCTGAAGCCGATGATCCAGTGGCTGAA 827
266 PheGlnCysLysValArgSerAspValLysProValIleGlnTrpLeuL 282
828 GCGGTGAGTAGCGCGCCGAGGGCGCCGACACATCCACATCGATGTGG 877
282 sarValAlaGlyTyrGlyAlaGlnGlyArgHisAsnSerThrIleAspValG 299
878 GCGGCGAAGATTGTGGTGTGCCACAGGGTGAAGCTGTGGTGGCGGGCC 927
299 lYgLyGlnLysPheValValLeuProThrGlyAspValTrpSerArgPro 315
928 GACGGCTCTACTCATATAGCTGTCTATCACCGCTGCCCGCCAGAGACA 977
316 AspGlySerTyrLeuAsnLysLeuLeuIleThrArgAlaArgGlnAspAs 332

```

```

978 TCGGGGCATGTACATCTGCTTGGCGCCACACACATGGGCTTACAGCTTCC 1027
332 pAlaGlyMetLysIleCysLysGlnValAlaAsnThrMetGlyTyrSerPheA 349
1028 GCAGGGCTTCTCAGCGTGTGGTGGCCAGACCCAAACCGCCAGGGGACACT 1077
349 rSerAlaPheLeuThrValLeuProAspProLysProGlnGlyProPro 365
1078 GTGGCTCTCTGCTCGGCCACTAGCCTGCGGTGGCCCGGTGATTCAGG 1127
366 ValAlaSerSerSerSerAlaThrSerLeuProTrpProValValIleG 382
1128 CATCCAGCCGCGCGCTGCTTTCATCTGTGGACACCTGCTGTGGCTTT 1177
382 yLeProAlaGlyAlaValPheIleLeuGlyThrLeuLeuLeuTrpLeuC 399
1178 GCCAGGCCAGAGAGAGCGGTGACCGCCCGCGCGCTGCCCTGCCCTGCT 1227
399 ySGlnAlaGlnLysLysProCysThrProAlaProAlaProProLeuPro 415
1228 GGGACCGCCCGCGCGGAGCGCCCGCCAGCCGACGAGGAGACAAGACCT 1277
416 GlyHisArgProProGlyThrAlaLeuAspArgSerGlyAspLysAspLe 432
1278 TCCTCGTGGCGCGCCCTCAGCGCTGGCGCTGTGTGTGTGAGG 1327
432 uProSerLeuAlaAlaLeuSerAlaGlyProGlyValGlyLeu..... 446
1328 AGCATGGGTCTCCGGACAGCCCGCCAGCCTTACTGGGCGCCAGGCCAATT 1377
446 ..... 446
1378 GCTGGCCCTAAGTTGTACCCCAACTCTACAGACATCCACACACAC 1427
446 ..... 446
1428 ACACGTGATTGCGCGCGCTGTGTGAGAGCATGGGTCCGGGACGCC 1477
447 .....Cys..GlnGlnHisGlySerProAlaAlaPr 456
1478 CCAGACATTAATGGGCGCCAGCGCCAGTGTGGCGCTGAAGTTGATCCCA 1527
456 oGlnHisLeuLeuGlyProGlyProValAlaGlyProLysLeuTyrProL 473
1528 AACTCTACAGACATCCACACACACACACACACACTCTCAGACACAC 1577
473 yLeuTyrThrAspIleHisThrHisThrHisThrHisSerHisThrHis 489
1578 TCACACGTGAGAGGCGAAGTCCACACAGACATCCACTATCAGTGC 1622
490 SerHisValGlnGlnLysValHisGlnHisIleHisTyrGlnCys 504
seq_name: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT.AAB19114
seq_documentation_block:
ID AAB19114 standard; Protein: 529 AA.
AC
XX AAB19114;
XX
DT 19-FEB-2001 (first entry)
XX
DE Polypeptide isolated from lymph node stromal cells of fsn -/- mice.
XX
KW Lymph node stromal cell; fsn -/- mice; inflammatory disorder;
KW immune system disorder; cancer; viral infection;
KW blood vessel growth; tumor necrosis factor disorder; arthritis;
KW inflammatory bowel disease; fibroblast growth factor-mediated disorder;
KW cardiac failure.
OS
XX Mus sp.
XX
PN W0200058463-A1.

```

XX 05-OCT-2000.  
 PD 18-FEB-2000: 2000MO-N200015.  
 PF  
 XX  
 PR 25-MAR-1999: 99US-0276268.  
 XX 26-AUG-1999: 99US-0383586.  
 PA (GENE-) GENESIS RES & DEV CORP LTD.  
 XX  
 PI Strichan L, Sleeman M, Abernethy N, Onrust R, Kumble KD;  
 PI Murison JG;  
 DR WPI: 2000-664924/64.  
 XX N-PSDB: AAA6724.  
 PT Polypeptide expressed in mammalian fsn +/- lymph node stromal cells,  
 PT useful for modulating growth of blood cells, for treating inflammatory  
 PT and tumour necrosis factor-mediated disorders, cancer and viral  
 PT disorders -  
 XX  
 PS Claim 1: Page 51-52; 75pp; English.  
 XX

CC The present sequence represents a polypeptide sequence which is  
 CC isolated from lymph node stromal cells of fsh +/- mice. The  
 CC polynucleotides and their polypeptides are useful for treating an  
 CC inflammatory disorder, disorder of immune system and cancer selected  
 CC from epithelial, lymphoid, myeloid, stromal and neuronal cancers, a  
 CC viral disorder, in particular HIV infection and for modulating the  
 CC growth of blood vessels. The polypeptides are useful for treating a  
 CC tumour necrosis factor (TNF) mediated disorder, such as those selected  
 CC from arthritis, inflammatory bowel disease and cardiac failure and a  
 CC fibroblast growth factor-mediated disorder. It is also useful in assays  
 CC to determine biological activity, to raise antibodies, to isolate  
 CC corresponding ligands or receptors, to quantify levels of protein or  
 CC cognate corresponding ligand or receptors, as anti-inflammatory agents,  
 CC and in compositions for the treatment of skin, connective tissue and  
 CC immune system diseases. The polynucleotide is useful as marker for  
 CC tissue, as a chromosome marker or tags in the identification of a  
 CC genetic disorder.  
 CC  
 XX

SO Sequence 529 AA:

alignment\_scores:  
 Quality: 2199.00 Length: 542  
 Ratio: 4.749 Gaps: 8  
 Percent Similarity: 85.424 Percent Identity: 78.413

alignment\_block:  
 US-09-598-042A-2 x AAB19114 ..

Align seg 1/1 to: AAB19114 from: 1 to: 529

94 CGGAGGTGGCCCGCGCTGGCCGACATATGGGTGACATGCCAGTGA 143  
 |||||||  
 34 ArgInvalAlaArgLeuGlyArgThrValArgLeuGlnCysProValGI 50  
 144 GGGGACCCGCGCGCTGACATGTGACCAAGAGATGCCGACCATCC 193  
 |||||||  
 50 uGlyAspProProleuthrmetrPhnrlYsAspGlyArgThrIleH 67  
 194 ACAGCGCTGAGCGCTTCCGCGTGTGCGCCAGGAGGGCTGAAGTGAAG 243  
 |||||||  
 67 IsSerGlyTTPSerArgPheArgValLeuProGlnGlyLeuLysValIys 83  
 244 CAGGTGAGGGGAGATGGCGGCTGAGTGTGCAAGGCCACCAACGG 293  
 ::::|  
 84 GluValAluIuagIuaSpAlaGlyValIryValCysLysAlaThrnsngI 100  
 294 CTTGCGGAGCGCTGACGCTCAACTACACCGCTGTCGCTGGATGACATTA 343  
 |||||||  
 100 yPheGlySerLeuSerValAsnYrThrIleuIleIleMetAspAspIles 117

344 GCCAGGGAAGAGAGCCTGGGCGCGACAGCTCTCTGGGGGTCAAGAG 393  
 |||||||  
 117 erProGlyLysGluSerProGlyProGlyLysIleSerGlyGlyGln 133  
 394 GACCCCGCAGCCAGCAGTGGGCGACGCGCGCTTCAACAGCCCTCCAA 443  
 |||||||  
 134 AspProAlaSerGlnGlnIrrAlaArgProArgPheThrGlnProSerIly 150  
 444 GATGAGGCGCGGGGTGATGCGAGGCGCGGTGGTACGCTCGCGGCTCA 493  
 |||||||  
 150 sMetArgAlaArgValIleAlaArgProValGlySerSerValArgLeuL 167  
 494 AGTGGTGGCCAGCGGCGACCTCGGCGCGACATCACTGATGAAGGAG 543  
 |||||||  
 167 yScysValAlaSerGlyHisProArgProAspIleMetTrpMetLysAsp 183  
 544 GACCGAGCCTTGACCGCGCCAGAGCGCGCTGAGCCAGCAAGAGAAGTG 593  
 |||||||  
 184 AspGlnThrLeuThrHisLeuGlnAlaSerGlnHisArgLysLysIstr 200  
 594 GACACTGAGCTGAAGACCTGGCGCGGAGGAGACGGCAATACACT 643  
 |||||||  
 200 pThrLeuSerLeuLysAsnLeuLysProGluAspSerGlyLysYrThrC 217  
 644 GCGCGGTGTGCAACCGCGCGCGCCATCAACGCCACTACAGGTGAT 693  
 |||||||  
 217 ySarGlyAlaSerAsnLysAlaGlyAlaIleAsnAlaThrYrLysValAsp 233  
 694 GTATCCAGCGGACCGCTTCCAAAGCCGCTGTCAAGAGCAGCAGCCCGT 743  
 |||||||  
 234 ValIleGlnArgThrArgSerLysProValLeuThrGlyThrHisProVa 250  
 744 GAACAGAGCGGTGAGCTGGCGGAGACAGCTCTCCAGTCAAGAGTGC 793  
 |||||||  
 250 IAsnThrThrValAspPheGlyLysThrThrSerPheGlnCysLysValA 267  
 794 GCAGCGACGTGAAGCCGCTGATCAGTGGCTGAAGCGCGTGAAGTACGG 843  
 |||||||  
 267 rGSerAspValLysProValIleGlnTrpLeuLysArgValGlyYrGly 283  
 844 GCGGAGGCGCGCCACAACCTCAACATCGATGTGGCGCGCAGAAGTTGT 893  
 ::::|  
 284 SerGlnGlyArgHisAsnSerThrIleAspValGlyLysPheVa 300  
 894 GTGCTGCGCCAGCGGTGACGTGTGTCGCGCGCCAGCGCTCTACTCA 943  
 |||||||  
 300 lValLeuProThrGlyAspValIrrPheSerArgProAspGlySerTyrLeuA 317  
 944 ATAAGCTGCTATCAACCGCTGCGCGCAGAGAGCATGCGGCGATGTACATC 993  
 |||||||  
 317 snLysLeuLeuIleSerArgAlaArgGlnAspAspAlaGlyMetYrIle 333  
 994 TGGCTGGGCGCCAAACCAATGGGCTACAGTTCGCGAGCGCTCTCTCAC 1043  
 |||||||  
 334 CysLeuGlyAlaAsnThrMetGlyTyrSerPheArgSerAlaPheLeuTh 350  
 1044 CGTGTGCGAGACCAAAACCGCGCAGGCGCACCTGTGGCGCTCTCTCTCT 1093  
 |||||||  
 350 rValLeuProAspProLysProProGlyProPrometAlaSerSerSers 367  
 1094 CGGCACTAGCCTGCGCGCGCGGTGCTGATCGGATGCCAGCGCGCGCT 1143  
 ||:::|  
 367 erSerThrSerLeuProTrpProValValIleGlyIleProAlaGlyAla 383  
 1144 GTTTCATCTGGGCAACCGCTGCTGCTGTGGCTTGGCAGGCCCAAGAA 1193  
 |||||||  
 384 ValPheIleLeuGlyThrValLeuLeuTrpLeuCysGlnThrLysLysLy 400  
 1194 GCGGTGACCCCGCGCGCTGCGCGCTGCGCTGCGCGACCGCGCGCG 1243  
 |||||||  
 400 sProCysAlaProAlaSerThrLeuProValProGlyHisArgProProG 417

```

1244 GGACGGCCGGACGGACGGAGACAAGACCTTCCTGTTGGCGGCC 1293
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
417 lYrHrSerArgGluArgSerGlyAspLysAspLeuProSerLeuAla... 432
1294 CTCAGCGCTGGCCCTGGTGTGGGGCTGTGAGAGACATGGGTCTCCGCG 1343
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
433 .....ValGlyIleCysGlnGluIuhIsGlySerAlaMe 443
1344 AGCCCCCAGACACTTACTGGGCCCGACCAATGCTGGCCCTAAGTTGT 1393
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
443 tAlaPrGlnHisIleLeuAlaSerGlySerThrAlaGlyProLysLeuT 460
1394 ACCCAAACTCTACACAGACATCCACACACACACACACTGATATGGCGCC 1443
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
460 yRProLysLeuTyrThrAspValHis..... 468
1444 GCCTGTGTGAGAGACATGGGTCTCCGCGACGCCCGACACTTACTGGGC 1493
468 ..... 468
1494 CCAGGCCCACTGCTGGCCCTTAAGTTTACCACCAACTCTACACACACAT 1543
469 .....T 469
1544 CCACACACACACACACACACTCTACACACACTCAGCGTGGAGGCA 1593
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
469 hr.HisThrHisThrHisThrCysThrHisThrLeuSerCysGlyGlyG 485
1594 AGGT.....CCACACACACATCCACTATCATGCTCTAGACGCGCAT 1637
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
485 nGlySerSerThrProAlaCysProLeuSerValLeuAsn...ThrAla 501
1638 CTGCAGAG.....GGCAGCGGGGGCGCGCCGACACAG 1669
:::|||||:|||||:|||||:|||||:|||||:|||||:
501 nLeuGlnAlaLeuCysProGluValGlyIleTyrGlyPro...ArgGln 516
1670 GCACACTGGGAGATGGAGGAC 1691
517 .GlnValGlyArgIleGln 523

seq_name: /sids1/gcgdata/geneseq/geneseqp/AA2000.DAT.AAB19115
seq_documentation_block:
ID AAB19115 standard: protein: 439 AA.
XX AAB19115:
XX AC
XX DT 19-FEB-2001 (first entry)
XX DE Polypeptide isolated from lymph node stromal cells of fsn -/- mice.
XX KW Lymph node stromal cell; fsn -/- mice; inflammatory disorder;
XX KW immune system disorder; cancer; viral disorder; HIV infection;
XX KW blood vessel growth; tumour necrosis factor disorder; arthritis;
XX KW inflammatory bowel disease; fibroblast growth factor-mediated disorder;
XX KW cardiac failure.
XX OS Mus sp.
XX PN WO200058463-A1.
XX PD 05-OCT-2000.
XX PF 18-FEB-2000; 2000WO-N200015.
XX PR 25-MAR-1999; 99US-0276268.
XX PR 26-AUG-1999; 99US-0383586.
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX PA Strachan L, Sleeman M, Abernethy N, Onrust R, Kumble KD;
XX PI Murlison JG;
XX

```

```

DR WPI: 2000-664924/64.
DR N-PSDB: AAN6725.
XX
XX Polypeptide expressed in mammalian fsn -/- lymph node stromal cells,
XX PT useful for modulating growth of blood cells, for treating inflammatory
XX PT and tumour necrosis factor-mediated disorders, cancer and viral
XX PT disorders
XX
XX Claim 1: Page 52-53; 75pp; English.
PS
XX
XX The present sequence represents a polypeptide sequence which is
XX CC isolated from lymph node stromal cells of fsn -/- mice. The
XX CC polynucleotides and their polypeptides are useful for treating an
XX CC inflammatory disorder, disorder of immune system and cancer selected
XX CC from epithelial, lymphoid, myeloid, stromal and neuronal cancers, a
XX CC viral disorder, in particular HIV infection and for modulating the
XX CC growth of blood vessels. The polypeptides are useful for treating a
XX CC tumour necrosis factor (TNF) mediated disorder, such as those selected
XX CC from arthritis, inflammatory bowel disease and cardiac failure and a
XX CC fibroblast growth factor-mediated disorder. It is also useful in assays
XX CC to determine biological activity, to raise antibodies, to isolate
XX CC corresponding ligands or receptors, to quantify levels of protein or
XX CC cognate corresponding ligand or receptors, as antiinflammatory agents,
XX CC and in compositions for the treatment of skin, connective tissue and
XX CC immune system diseases. The polynucleotide is useful as marker for
XX CC tissue, as a chromosome marker or tags in the identification of a
XX CC genetic disorder.
XX
XX
XX Sequence 439 AA:
S0
XX
XX alignment_scores:
XX Quality: 1779.50 Length: 463
XX Ratio: 4.634 Gaps: 9
XX Percent Similarity: 82.937 Percent Identity: 75.810
XX
XX alignment_block:
XX US-09-598-042a-2 x AAB19115
XX
XX Align seg 1/1 to: AAB19115 from: 1 to: 439
XX
XX 334 GATGACATTAGCCCGAGGAGAGAGCCCTGGGCCCGACAGCTCTCTGG 383
XX |||||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 23 AspAspIleSerProGlyLysGlnSerProGlyLysSerSerG 39
XX 384 GGATCAAGAGAGCCCGCCAGCAGAGTGGGACAGACCGGCTTCACAC 433
XX |||||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 39 yGlyGlnGlnAspProAlaSerGlnGlnTyrAlaArgProAlaPheThrG 56
XX 434 AGCCCTCCAGATAGAGCGCGGGGTGATCGACGCGCCGTTGAGTCC 483
XX |||||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 56 InProSerLysMetArgArgValIleAlaArgProValGlySerSer 72
XX 484 GTGGGGCTCAAGTGGCGGAGCGGCGACCCCTGGCCCGCATCAGGTG 533
XX |||||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 73 ValArgLeuLysCysValAlaSerGlyHisProArgProAspIleMetTr 89
XX 534 GATGAGAGAGAGAGCGCTTGAGCGCGCCAGAGCGCGCTGAGCCAGGA 583
XX |||||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 89 pMetLysAspAspGlnThrIleuThrHisLeuGlnAlaSerGlnHisArgL 106
XX 584 AGAAGAGTGGACACTGAGCTGAAGAACTGCGCGCGGAGAGACGCGC 633
XX |||||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 106 yslYslYslYslrPrThrLeuSerLeuLysAsnLeuLysProGlnAspSerGly 122
XX 634 AAATACACCTGCGCGTGTGACACCGCGGCGGCATCAACGCGCACCTA 683
XX |||||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 123 LysTyrThrCysArgValSerAsnLysAlaGlyAlaIleAsnAlaThrTy 139
XX 684 CAAGGTGATGTGATCCAGCGGACCGCTTCCAGCCCGTGTCTCAGAGCA 733
XX |||||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 139 rLysValAspValIleGlnArgThrArgSerLysProValLeuThrGlyT 156
XX

```

```

734 CGACCCCGTGGACAGACAGCGGTGAGCTTGGGGGGACCGCTCTCCAG 783
|||||
156 hrhIspVAlAsnThrThrValAspPheGlyGlyThrThrSerPheGln 172
|||||
784 TGCAAGGTGGCGCAGCAGCGTGAAGCGGTGATCCAGTGGCGAAGCCCGT 833
|||||
173 CysIysValAspSerAspValLysProValIleGlnThrLeuLysArgVa 189
|||||
834 GGAATAGCGGCGCGAGGCGCGCCAACTCCACATCATGATGTGGCGCGC 883
|||||
189 IgluTyGlySerIgluLysArgHisAsnSerThrIleAspValGlyGly 206
|||||
884 AGAAGTTGTGTGCTGCTCCACAGGCGTACGTGTGTGGCGCGCGAGGC 933
|||||
206 InLysPheValIleValLeuProThrGlyAspValIlePheArgProAspGly 222
|||||
934 TCCACTCATATAGTGTGCTCATACCGCGTGGCGCGCGAGACGATGGCGG 983
|||||
223 SerTyLeuAsnLysLeuIleLeuSerArgAlaArgGlnAspAlaGly 239
|||||
984 CATGTACATGTGCTTGGCGCGCAACACATGGCTACAGCTTCCGACAGC 1033
|||||
239 YMetTyIleCysLeuGlyAlaAsnThrMetGlyTyrSerPheArgSerA 256
|||||
1034 CCTTCCACCGCTGCTGCCAGACCCAAA...CCGCCAGGCGCACCTGTG 1080
|||||
256 IapheLeuThrValLeuProAspProLysProProGlyProPromet 272
|||||
1081 GCCCTCTGCTCTGCGCAGCTAGCGCTGCGCGCGTGGCGCATGGCAT 1130
|||||
273 AlaSerSerSerSerThrThrSerLeuProThrProValIleGlyIle 289
|||||
1131 CCAGACCGCGCGCTGTCTTTCATCTGGGACCGCTGCTGCTGCTTGGC 1180
|||||
289 eProAlaGlyAlaValPheIleLeuGlyThrValLeuLeuTrpLeuGly 306
|||||
1181 AGGCGCCAGAGAGCGGTGACCCCGCGCGCTGCGCTGCGCTGCGCTGG 1230
|||||
306 InThrLysLysLysProCysAlaProAlaSerThrLeuProValProGly 322
|||||
1231 CACCGCGCGCGCGGAGCGCGCGCGACCGCAGCGAGAGACAGACCTTC 1280
|||||
323 HisArgProProGlyThrSerArgGluArgSerGlyAspLysAspLeuP 339
|||||
1281 CTGTTGGCGCGCTTACGCGCTGCGCTGTGGGGCTGTGAGAGAGC 1330
|||||
339 oSerLeuAla.....ValGlyIleCysGluGluH 349
|||||
1331 ATGGGTCTCCGGGACCGCGCGCACTTACTGGGCGCGCGCGAGTGTCT 1380
|||||
349 IsGlySerAlaMetAlaProGlnHisIleLeuAlaSerGlySerThrAla 365
|||||
1381 GGGCTTAAGTTGTAACCAACTCTACACAGACATCCACACACACACACA 1430
|||||
366 GlyProLysLeuTyProLysLeuTyThrAspValHis..... 378
|||||
1431 CTGATTTGGCGCGCGCTGTGTAGAGAGCATGGGTCTCCGACCGCCCA 1480
|||||
378 ..... 378
1481 GCACTTACTGGGCGCGCGCGAGTGTGCTGCGCTAAGTTGTACCCCAAC 1530
|||||
378 ..... 378
1531 TCTACACAGACATCCACACACACACACACACTCTACACACACTCA 1580
|||||
379 .....Thr.HisThrHisThrHisThrCysThrHisThrLeu 391
|||||
1581 CAGGTGAGGGGCAAGT.....CCACGACACATGCACATATGAGTCTA 1624
|||||
391 eCysGlyGlyGlyGlySerSerThrProAlaCysProLeuSerValLeu 407
|||||
1625 GACGCGACCGTATCTGCAGAG.....GCGACGCGGGGG 1656

```

```

seq_name: /sids1/gcgdata/geneseq/geneseq/AA2000.DAT:AA19116
seq_documentation_block:
ID   AA19116 standard; Protein; 322 AA.
XX
XX   AAB19116;
XX
XX   19-FEB-2001 (first entry)
XX
DE   Homologue of polypeptide from lymph node stromal cells of fsn -/- mice.
XX
XX   Lymph node stromal cell; fsn -/- mice; inflammatory disorder;
XX   immune system disorder; cancer; viral infection; HIV infection;
XX   blood vessel growth; tumour necrosis factor disorder; arthritis;
XX   inflammatory bowel disease; fibroblast growth factor-mediated disorder;
XX   cardiac failure.
XX
XX   Homo sapiens.
XX
XX   WO20058463-A1.
XX
XX   05-OCT-2000.
XX
XX   18-FEB-2000; 2000WO-NZ00015.
XX
XX   25-MAR-1999; 99US-0276268.
XX
XX   26-AUG-1999; 99US-0383586.
XX
XX   (GENE-) GENESIS RES & DEV CORP LTD.
XX
XX   Strachan L, Sleeman M, Abernethy N, Onrust R, Kumble KD;
XX   Murison JG;
XX
XX   WPI: 2000-664924/64.
XX
XX   N-PSDB: AAA96726.
XX
XX   Polypeptide expressed in mammalian fsn -/- lymph node stromal cells,
XX   useful for modulating growth of blood cells, for treating inflammatory
XX   PT and tumour necrosis factor-mediated disorders, cancer and viral
XX   PT disorders
XX
PS   Claim 1; Page 53-54; 75pp; English.
XX
XX   The present sequence represents a homologue of a polypeptide sequence
XX   which is isolated from lymph node stromal cells of fsn -/- mice. The
XX   CC polynucleotides and their polypeptides are useful for treating an
XX   CC inflammatory disorder, disorder of immune system and cancer selected
XX   CC from epithelial, lymphoid, myeloid, stromal and neuronal cancers, a
XX   CC viral disorder, in particular HIV infection and for modulating the
XX   CC growth of blood vessels. The polypeptides are useful for treating a
XX   CC tumour necrosis factor (TNF) mediated disorder, such as those selected
XX   CC from arthritis, inflammatory bowel disease and cardiac failure and a
XX   CC fibroblast growth factor-mediated disorder. It is also useful in assays
XX   CC to determine biological activity, to raise antibodies, to isolate
XX   CC corresponding ligands or receptors, to quantify levels of protein or
XX   CC cognate corresponding ligand or receptors, as antiinflammatory agents,
XX   CC and in compositions for the treatment of skin, connective tissue and
XX   CC immune system diseases. The polynucleotide is useful as marker for
XX   CC tissue, as a chromosome marker or tags in the identification of a
XX   CC genetic disorder.
XX
SO   Sequence 322 AA;

```

```

alignment_scores:
Quality: 1497.00      length: 317

```







```

1045 GTGCTGCCAGACCCAAACCCGACGGCCACCTGTGCTCTCTCTCTC 1094
|||||
52 ValLeuProAspProLyProProGlyProPrometAlaSerSerSe 68
1095 GGCACACTAGCTCCGCTGGCCGCTGTGATCGGACATCCACGGCGCTG 1144
|||||
68 rSerThrSerLeuProTrProValIleGlyLeProAlaGlyAlav 85
1145 TCTTCATCTCTGGGACACCTCTCTCTGTGCTTTTGCCAGGCCAGAGAG 1194
|||||
85 aIPheIleLeuGlyThrValLeuLeuTrpLeuGlyGlnThrLysLys 101
1195 CCGTGACACCCCGCGCTCCCTCCCTGCTGGGACCGCGCGCGG 1244
|||||
102 ProCysAlaProAlaSerThrLeuProValProGlyHisArgProProG 118
1245 GACGGCCCGGACCGCGACGGAGACAGACCTTCCCTGTTGGCGCGCC 1294
|||||
118 YThrSerArgGlnArgSerGlyAspLysAspLeuProSerLeuAla... 133
1295 TCAGCGCTGCGCCCTGGTGGGCTGTGTGAGAGACATGGGTCTCCGCA 1344
|||||
134 .....ValGlyIleCysGlnGlnHisGlySerAlaMet 144
1345 GCGCCCGACACTTACGTGGGCGGCGGCGGAGTGTGCTGAGCTTGTGTA 1394
|||||
145 AlaProGlnHisIleLeuAlaSerGlySerThrAlaGlyProLysLeuTy 161
1395 CCGCAAACTTACACAGACATCCACACACACACACTGTATTGGCGCGG 1444
|||||
161 rProLysLeuTyThrAspValHisThrHis..... 173
1445 CCTGTGTGAGAGACATGGGCTCTCCGGACGCCCGACGACTTACTGGCC 1494
|||||
173 ..... 173
1495 CAGGCGCAGTTGTGGCCCTAAGTTGTACCCCAAACTGTACACAGACATC 1544
|||||
173 ..... 173
1545 CACACACACACACACACACTCTCACACACACTCACAGTGGAGGCGCA 1594
|||||
174 ThrHisThrCysThrHisThrLeuSer.....CysTrpArgAla 187
1595 GG.....TCCACCGACGACATCCACTATCATGCTAGA 1626
|||||
187 rGpHeIleAsnThrSerMetSerThrIleSerAlaLys 199
seq_name: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:AA19113
seq_documentation_block:
ID AAB19113 standard; Protein: 126 AA.
XX
AC AAB19113:
XX
DT 19-FEB-2001 (first entry)
XX
DE Polypeptide isolated from lymph node stromal cells of fsn -/- mice.
XX
KW Lymph node stromal cell; fsn -/- mice; inflammatory disorder;
KW immune system disorder; cancer; viral infection; HIV infection;
KW blood vessel growth; tumour necrosis factor disorder; arthritis;
KW inflammatory bowel disease; fibroblast growth factor-mediated disorder;
KW cardiac failure.
XX
OS Mus sp.
XX
FT Key location/Qualifiers
XX
FT Misc-difference 123 /note="unknown amino acid encoded by GNGATC"
XX
XX WO200058463-A1.

```

```

PD 05-OCT-2000.
XX
PF 18-FEB-2000; 2000WO-NZ00015.
XX
PR 25-MAR-1999; 9905-0276263.
XX
PR 26-AUG-1999; 9905-0383585.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Strachan L, Sleeman M, Abernethy N, Onrust R, Kumble KD;
PI Morrison JG;
XX
DR WPI: 2000-664924/64.
XX
DR N-PSDB; AAA96723.
XX
PT Polypeptide expressed in mammalian fsn -/- lymph node stromal cells,
PT useful for modulating growth of blood cells, for treating inflammatory
PT and tumour necrosis factor-mediated disorders, cancer and viral
PT disorders -
XX
PS Claim 1; Page 50-51; 75pp; English.
XX
CC The present sequence represents a polypeptide sequence which is
CC isolated from lymph node stromal cells of fsn -/- mice. The
CC polynucleotides and their polypeptides are useful for treating an
CC inflammatory disorder, disorder of immune system and cancer selected
CC from epithelial, lymphoid, myeloid, stromal and neuronal cancers, a
CC viral disorder, in particular HIV infection and for modulating the
CC growth of blood vessels. The polypeptides are useful for treating a
CC tumour necrosis factor (TNF) mediated disorder, such as those selected
CC from arthritis, inflammatory bowel disease and cardiac failure and a
CC fibroblast growth factor-mediated disorder. It is also useful in assays
CC to determine biological activity, to raise antibodies, to isolate
CC corresponding ligands or receptors, to quantify levels of protein or
CC cognate corresponding ligand or receptors, as anti-inflammatory agents,
CC and in compositions for the treatment of skin, connective tissue and
CC immune system diseases. The polynucleotide is useful as marker for
CC tissue, as a chromosome marker or tags in the identification of a
CC genetic disorder.
XX
SQ Sequence 126 AA;
XX
alignment_scores:
Quality: 645.50 Length: 127
Ratio: 5.164 Gaps: 1
Percent Similarity: 98.425 Percent Identity: 95.276
alignment_block:
US-09-598-042A-2 x AAB19113 ..
Align seg 1/1 to: AAB19113 from: 1 to: 126
754 GTGACTTCGGGGGAGCACCGTCTTCAGTGCAGAGTGGCAGGACGT 803
|||||
1 ValAspPheGlyGlyThrThrSerPheGlnCysLysValArgSerAspVa 17
804 GAAGCGGTGATCCAGTGGCTGAAGCGGCTGGAGTAGAGCGCGCGAGGCC 853
|||||
17 llyseroValIleGlnTrpLeuLysArgValGlnTrpGlySerGlnGly 34
854 GCCACACTCCACCATCATGTGGGCGGCGCAGAACTTTGTGTGGCC 903
|||||
34 rGHisAsnSerThrIleAspValGlyGlyGlnLysPheValValLeuPro 50
904 ACGGTGACGTGTGTGCGGCGCGGCGGCTCTCTCAATAAAGCTCT 953
|||||
51 ThrGlyAspValTrpSerArgProAspLysSerTyTrpLeuAsnLysLeu 67
954 CATCACCCGTCGCCCGCGACGAGATGCGGCGCATGATCTGCTTGGCG 1003
|||||
67 uIleSerArgAlaArgGlnAspAspAlaGlyMetTyTrpIleCysLeuGly 84

```

```

1004 CCAACACCATGGCTACAGCTTCGCGAGCGGCTCTACCGTGCTGCCA 1053
|||||
84 laasThnMetGlyTyrSerPheArgSerAlaPheLeuThrValLeuPro 100
|||||
1054 GACCCAAACCGCCAGGCGCACCTGTGCTCCCTCCCTCGGCGCATAG 1103
|||||
101 AspProLysProProGlyProPheMetAlaSerSerSerSerThrSe 117
|||||
1104 CCTGCGCTGGCGGCTGTCATCGGCATGCCA 1134
|||||
117 rLeuProTyrPro...Val**GlyIlePro 126
|||||

seq_name: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT:AA57131
seq_documentation_block:
ID AAB57131 standard; Protein; 101 AA.
XX
AC AAB57131:
XX
DT 13-MAR-2001 (first entry)
XX
DE Human prostate cancer antigen protein sequence SEQ ID NO:1709.
XX
KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;
KW vulnerrary; gastrointestinal; nephrotoxic; antineoplastic; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease.
XX
OS Homo sapiens.
XX
PN WO200055174-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05988.
XX
PR 12-MAR-1999; 990S-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI: 2000-587513/55.
XX
DR N-PSDB; AAF16334.
XX
PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
XX
PS Claim 11; Page 2183; 2338pp; English.
XX
CC AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytosolic,
CC cardioactive, immunomodulatory, muscular, vulnerrary, gastrointestinal,
CC nephrotoxic, antineoplastic, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF15506 to AAF15514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 101 AA:

```

```

Quality: 529.00 Length: 100
Ratio: 5.290 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.000

alignment_block:
US-09-598-042a-2 x AAB57131 ..
Align seg 1/1 to: AAB57131 from: 1 to: 101

499 GTGGCCAGCGGGACCGCTGGCCGACATCAGTGGATGTAAGAGACGCA 548
|||||
1 ValAlaSerIleHisProArgProAspIleThrTyrMetLysAspAspG1 17
|||||
549 GGCCCTGACGCGGCCAGAGCGCGCTGAGCCAGCAAGAAAGACTGACAC 598
|||||
17 nAlaLeuThrArgProGluAlaIleGluProAlrGlyLysLysTyrThrL 34
|||||
599 TGAGCCGAAAGAACTGCGCGCGGAGAGACAGCGCAATACCTGCCGC 648
|||||
34 euserLeuLysAsnLeuAlrGProGluAspSerGlyLysTyrThrCysArg 50
|||||
649 GTGTGCAACCGCGCGGCGCATCAAGCCACCTACAGGTGATGTGAT 698
|||||
51 ValSerAsnArgAlaGlyAlaIleAsnAlaThrTyrLysValAspVal11 67
|||||
699 CCAAGCGACCGCTTCCAAAGCGCGTCTCACAGGACGACGCCCGTGACA 748
|||||
67 eGlnArgThrArgSerLysProValLeuThrGlyThrHisProValAsn 84
|||||
749 CGAGCGTGACCTTCGGGGGAGACGACGCTTCGAGCAAGTGCGCCAGC 798
|||||
84 hrThrValAspPheGlyGlyThrThrSerPheGlnCysLysValAlrGThr 100
|||||

seq_name: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT:AA21080
seq_documentation_block:
ID AAR21080 standard; Protein; 821 AA.
XX
AC AAR21080;
XX
DT 20-MAY-1992 (first entry)
XX
DE flg receptor protein.
XX
KW Fibroblast growth factor receptor; heparin binding proteins;
KW tyrosine kinase; bacterially expressed kinase; CSF-1; PDGF.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Peptide 1..21
FT /label= signal_peptide
FT Protein 22..821
FT /label= bek_receptor_protein

W09200999-A.
XX
23-JAN-1992.
XX
03-JUL-1991; 91WO-US04745.
XX
PR 06-JUL-1990; 90US-0549587.
XX
PA (RORE ) RORER INT HOLDINGS.
XX
PI Dione CA, Crumley G, Jaye MC, Schlessinger J;
XX
DR WPI: 1992-056827/07.
XX
DR N-PSDB; AAQ21003.
XX
PT New fibroblast growth factor receptor proteins - useful in
PT treating GF-mediated conditions e.g. angiogenesis of tumours,
PT mitogenic effects in psoriasis, arthritis

```



```

XX Homo sapiens.
OS
XX MO9117183-A.
XX
XX 14-NOV-1991.
XX
XX
XX 25-APR-1991; 91MO-JP00557.
XX
XX 28-DEC-1990; 90JP-0415801.
XX
XX 27-APR-1990; 90JP-0113146.
XX
XX 31-JUL-1990; 90JP-0204438.
XX
XX 14-SEP-1990; 90JP-0245256.
XX
XX (TAKE ) TAKEDA CHEMICAL IND KK.
XX
XX Igarashi K, Senoo M, Watanabe T;
XX
XX WPI: 1991-353723/48.
XX
XX N-PSDB; AA014849.
XX
XX
XX New muten(s) of proteins - with fibroblast growth factor
XX receptor activity, useful for treating multiple endocrine
XX neoplasia, prostatic hypertrophy, used for diagnosis
XX
XX Example 2; Fig 4; 88pp; English.
XX
XX A cDNA library prepared from human cancer cell line Kato III mRNA
XX was screened with an oligonucleotide corresponding to amino acids
XX 529-541 of chicken basic FGF receptor. Three positive clones were
XX obtained and were cloned into pUC118/119 to give pTB1227, pTB1228
XX and pTB1229. The longest clone was in pTB1229. The amino acid
XX sequence was deduced from the insert sequence.
XX
SO Sequence 643 AA;

alignment_scores:
  quality: 452.50      Length: 385
  Ratio: 1.926         Gaps: 10
  Percent Similarity: 61.039   Percent Identity: 30.909

alignment_block:
US-09-598-042a-2 x AAR15267 ..
Align seg 1/1 to: AAR15267 from: 1 to: 643

100 GTGGCCGGCGGCGCCGACATATCGGCGTGCAGTCCCGACATGAGGGGA 149
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51 ValAlaIaIaProGlyGluSerLeuGluValArgCysLeuLeuLys.... 65
150 CCCGCCCGCGCTGACATGTGACCAAGATGCGCGACCATCCACAGCG 199
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
66 AspAlaIaValIleSerTrpThrLysAspGly.....ValHisLeuG 80
200 GCTGAGCCCGCTTCGCGTCTG...CCGAGGGGCTGAAGGTGAAGCAG 246
    || ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
80 LysProAsnAsnArgThrValIleuIleGlyIuTrpLeuGlnIleLeuGly 96
247 GTGAGGGGGGAGTGGCGGCTGTAGTGTGACAGGCCACCAACGCGCTT 296
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
97 AlaThrProAlaGaspSerGlyLeuTrpAlaCysThrAlaSerArgThrVa 113
297 CGGAGCGCTGACGTCAACTACACCTGTCGTGCTGATGACATTAAGCC 346
    ||| ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
113 LaspSerGluThrTrpTrpTrpMetValAsnValThrAspAlaIleSer 130
347 CAGGGAAGAGAGCGCTGGGGCCGACAGCTCTCTGCGGGGTCAAGAGAC 396
    ||| ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
130 ergIlyAspAsp.....GluAspAspThrAspGlyAlaGluAsp 142
397 CCGCCAGCCAG.....CAGTGGGCAGAGCCGCGCTTCACACAGCC 437
    ::|||::: ||| ::|||:::

```

```

143 PheValSerGluAsnSerAsnAsnLysArgAlaProTrpTrpThrAsnTh 159
438 CTCACAGATGAGCGCGCGGTGATGCGACGCGCGGTGGGTAGTCCGTGC 487
    ::|||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
159 rGluLysMetGluLysArgLeuHisAlaValProAlaIaIaSerThrVal 176
488 GGTCAAGTGGTGGCCAGCGCGGACCCCTGCGCCGACATCAGCTGATG 537
    ::|||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
176 yspHeArgCysProAlaGlyGlyAsnProMetProThrMetArgTrpLeu 192
538 AAGACGACGACGCGCTGAGC.....CGCCAGAGCGCGGTGAGCC 578
    |||||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
193 LysAsnGlyLysGluPheLysGlnGlnHisArgIleGlyIleTrpLysVa 209
579 CAGGAGAGAAGTGAACACTGACCTGAAGACCTGACCGCGCGAGGACA 628
    |||||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
209 LArgAsnGlnHisTrpSerLeuIleMetGluSerValValProSerSPL 226
629 GCGCAAAATACACCTCGCGCGGTGCAACGCGCGGGCGCCATCAGCGCC 678
    ::|||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
226 ysgIlyAsnTrpThrCysValValGluAsnGluTrpGlySerIleAsnHis 242
679 ACCTACAAAGTGTGATGATCAGCGGACCGCTCCAGCCCGCTGCTCAC 728
    |||||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
243 ThrTrpHisLeuAspValValGluArgSerProHisArgProIleLeuG 259
729 AGGCACGACCCCGCTGAACACGACGTGAGTTCGGGGGAGCACAGTCT 778
    ::|||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
259 nAlaGlyLeuProAlaAsnAlaSerThrValValGlyIlyAspValGluP 276
779 TCCAGTCAAGTGTGCGGACGACGTGAGCGCGGTGATCAGTGGCTGAG 828
    |||||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
276 heValCysLysValIlySerAspAlaGlnProHisIleGlnTrpIleLys 292
829 CGCGTGGAG.....TACGGCGCGGAGCGCGCCCAACTC 863
    ::|||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
293 HisValGluLysAsnGlySerLysTrpGlyProAspGly..... 305
864 CACCATGATGTGGCGCGCGCAGAGTTTGTGTGCTCCCGCACGGGTAGC 913
    |||||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
306 .....LeuProTrpIleuLysV 311
914 TG.....TGTGCGGCGCGCGACGCGCTCTCACTCAATTAAGCTGTCATC 957
    |||||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
311 alleuLysHisSerGlyIleAsnSerSerAsnAlaGluValLeuAlaLeu 327
958 ACCCGTCCCGCCGACGACATGCGGCGATGATCATCTGCTGGCGCAA 1007
    ::|||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
328 PheAsnValThrGluAlaAspAlaGlyGluTrpIleCysLysValSerAs 344
1008 CACCATGGGTACAGCTTCCGACGCGCGCTCTCACTGCTGCGCAAGCC 1057
    |||||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
344 nTrpIleGlyGlnAlaAsnGlnSerIleAlaTrpLeuThrValLeuProLysG 361
1058 CAAAACCGCAAGGCGCACCTGTGCGCTCTGCTGCGCGCACTAGCTG 1107
    ::|||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
361 InglnAlaProGly.....ArgGluLysGluIleThrAlaSer 373
1108 CCGTGGCGCGGTGATCATGCGCATCCGACGCGCGCGCTCTCTCATCCGGG 1157
    ::|||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
374 ProAspTrpTrpLeuGluIleAlaIleTrpCysIleGlyValPheLeuIleAl 390
1158 CACCTGCTCTGCTGCTTGGCCAGGCGCGCAAGAAGACCTGACCCCGG 1207
    ::|||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
390 acYsmetValValThrValIleLeuLysArgMetLysAsnThrThrLysL 407
1208 CGCCT 1212
    |||
407 yspPro 408

```

seq\_name: /std1/gcgdata/geneseq/geneseq/AA2000.DAT:AA97170

seq\_documentation\_block:



```

791 TGCGCAGCAGTGAAGCCGGTATCCATGCTGAAGCCGTTGAGTAC 840
    || ||||| ||||| ||||| ||||| ||||| ||||| |||||
277 ALTYSerAspProGlnProHisIleGlnIlePheLeuYSHisIleGluVal 293
841 GCGCCGAGGCGCCGACACTCCACCATGATGGCGCCGAGAAATT 890
    ::::: ||| ||| :::::
294 .....AsnGlySerLeuIleGlyProAspAsnLeuProTyr 305
891 TGTG...GTGTCGCCACGAGGTGACGTGTGTCGCCGCGAGGCGTCT 937
    ||| ||||| ||||| ||||| ||||| ||||| |||||
305 rValGlnIleLeuYSHisIleGluVal.....AsnThrThraspIysG 320
938 ACSTCAATAACTGCTCATCAGCCGTCGCCGCGAGAGATGCGGCGATG 987
    ::::: ||| :::::
320 lMetGluValLeuHisIleuArgAsnValSerPheGluAspAlaGlyGlu 336
988 TACATCTGCTGCGCCGACACACATGAGTTCGCGAGCGCTT 1037
    ||| ||||| ||||| ||||| ||||| ||||| |||||
337 TyrThrCysLeuAlaGlyAsnSerIleGlyLeuSerHisIleSerAlaTr 353
1038 CCTCAGCGTGTGCGCAGCCGACAAAGCCGAGGCGCAGCTGTGCTGCT 1087
    ::||| ||||| ||||| ||||| ||||| ||||| |||||
353 PheuThrValLeu...GluAlaLeuGluGluArgProAlaValMetHrs 369
1088 GTCCTCGGCGCAGTACGCTGCGCTGCGCTGCTCATCGGCGATCCAGCC 1137
    || ||| ||| |||||
369 er.ProleuTyrLeuGluSerArgGly..... 377
1138 GCGCCTGTCTCATCTGGGCGACCTGTCTGTGCTTGCAGGCGCA 1187
    ||||| ||||| ||||| ||||| |||||
378 .....GlyLeuValProAlaGlySerGly.SerProGlyLeu 389
1188 GAGAGACCGCTGCACCCCGCGGCTGCGCTGCCCTGCTGGGAGCGCC 1237
    ::||| ||||| ||||| ||||| ||||| |||||
390 GlnGlu.....ProLysSerCysAspIysThrHisThrCysProPr 403
1238 GCGCGGGGAGCGCCCGGACCGCGAGCGAGACAGACCTTCCTC 1283
    ::::: ||| |||||
403 OCys.....ProAlaProGluLeuGlyGlyProSerVal 415
seq_name: /stid1/gcgdata/geneseq/geneseqp/AA1991.DAT:AA15268
seq_documentation_block:
ID AA15268 standard; Protein; 769 AA.
XX
AC AA15268;
XX
DT 18-FEB-1992 (first entry)
XX
DE Clone pTBI284-encoded complete FGF receptor.
XX
KW Human; fibroblast growth factor; cancer.
XX
OS Homo sapiens.
XX
PN MO9117183-A.
XX
PD 14-NOV-1991.
XX
PF 25-APR-1991; 91MO-JP00557.
XX
PR 28-DEC-1990; 90JP-0415801.
PR 27-APR-1990; 90JP-0113146.
PR 31-JUL-1990; 90JP-0204438.
PR 14-SEP-1990; 90JP-0245256.
XX
PA (TAKE ) TAKEDA CHEMICAL IND KK.
XX
PI Igarashi K, Senoo M, Watanabe T;
XX
DR WPI, 1991-353723/48.
DR N-PSDB; AAQ14268.
XX

```

```

PT New mutin(s) of proteins - with fibroblast growth factor
PT receptor activity, useful for treating multiple endocrine
PT neoplasia, prostatic hypertrophy, used for diagnosis
XX
XX Example 3; Fig 7; 88pp; English.
XX
CC A cDNA library prepared from human cancer cell line Kato III mRNA
CC was screened with an oligonucleotide corresponding to amino acids
CC 529-541 of chicken basic FGF receptor. Three positive clones were
CC obtained. One was cloned into pUC118/119 to give pTBI229 (see
CC AAQ14849). The complete FGF coding sequence was obtained by ligating
CC the insert from pTBI229 to the DNA sequence of the plasmid pTBI281
CC insert which encodes the carboxyl terminus of the FGF receptor from
CC Glu 533 onwards.
XX
SQ Sequence 769 AA;
XX
alignment_scores:
    Quality: 448.50      Length: 385
    Ratio: 1.909        Gaps: 10
    Percent Similarity: 61.039      Percent Identity: 30.649
alignment_block:
US-09-598-042a-2 x AA15268 ..
Align seg 1/1 to: AA15268 from: 1 to: 769
100 GTGGCGCGGCTGGGCGCGCACTATGCGGTGACATGCCAGTGGAGGGGGA 149
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
51 ValAlaAlaProGlyGluSerLeuGluValArgCysLeuLeuLys..... 65
150 CCGCGCGCGCGCTGACCATGTGACCAAGAGTGGCGCGACCATCCAGCGC 199
    ::::: ||||| ||||| ||||| ||||| |||||
66 .AspAlaAlaValIleSerThrPheIysAspGly.....ValHisLeuG 80
200 GCTGAGACCGCTTCGCGCTGCTG...CCGAGGCGGTGAAGTGAAGCAG 246
    ||| ::::: ||||| ||||| ||||| ||||| |||||
80 LysProAsnAsnArgThrValIleGlyLeuTyrLeuGlnIleLysGly 96
247 GTGAGCGGAGGAGATGCCGCGCTGTACGTGTGCAAGCCGACCAACGCTT 296
    ::||| ||||| ||||| ||||| ||||| ||||| |||||
97 AlaThrProArgAspSerGlyLeuTyrAlaCysThrAlaSerArgHrta 113
297 GCGCAGCGTGAAGCTCACTACACCTGCTGCTGATGACATTAGCC 346
    ||| ::::: ||||| ||||| ||||| ||||| |||||
113 LaspSerGluThrTyrPheMetValAsnValThrAspAlaIleSers 130
347 CAGGGAAGGAGAGCTGGGCGCGCAGAGCTCTGCGGGGTCAAGAGAGC 396
    ||| ::::: ||||| ||||| ||||| ||||| |||||
130 erGlyAspAsp.....GluAspAspThrAspGlyAlaGluAsp 142
397 CCGCGCAGCCAG.....CAGTGGGACAGACCGCGCTTCACACAGCC 437
    ::||| ||||| ||||| ||||| ||||| |||||
143 PheValSerGluAsnSerAsnAsnLysArgAlaProTyrThrHisnTh 159
438 CTCGAAGATGAGCGCGCGGTGATGCCAGCGCGCGGTAGCTCCGTCG 487
    ::||| ||||| ||||| ||||| ||||| ||||| |||||
159 rGluLysMetGluLysArgLeuHisAlaValProAlaAlaAsnThrVal 176
488 GGCCTCAAGTGCCTGCGCAGGCGACCTCGGCGCGACATCACTAGGATG 537
    ::||| ||||| ||||| ||||| ||||| ||||| |||||
176 yspHeArgCysProAlaGlyGlyAsnProMetProThrMetArgTyrLeu 192
538 AAGGAGACCAAGCGCTTGACG.....CGCCAGAGCGCGCTAGCC 578
    ||||| ::::: ||||| ||||| ||||| ||||| |||||
193 LysAsnGlyLysGluPheLysGlnGluHisArgIleGlyTyrTyrSva 209
579 CAGGAGAAGAAGTGAAGCACTGAGCTGAAGAAGCGCGCGGAGAGACA 628
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
209 LArgAsnGlnHisTrpSerLeuIleMetGluSerValValProSerGlu 226
629 GCGCAAAATACACTGCGCGGTGTGCAACCGCGCGGCGCATCAACGCC 678

```

```

PR 27-APR-1990: 90JP-0113146.
PR 31-JUL-1990: 90JP-0204438.
PR 14-SEP-1990: 90JP-0245256.
XX
XX
PA (TAKE ) TAKEDA CHEMICAL IND KK.
XX
PI Igarashi K, Senoo M, Watanabe T;
DR WPI; 1991-353723/48.
DR N-PSDB; AAQ14864.
XX
XX
PT New mutin(s) of proteins - with fibroblast growth factor
PT receptor activity, useful for treating multiple endocrine
PT neoplasia, prostatic hypertrophy, used for diagnosis
XX
XX
PS Claim 5; Fig 12; 88bp; English.
XX
CC A cDNA library prepared from human cancer cell line Kato III mRNA
CC was screened with an oligonucleotide corresponding to amino acids
CC 559-541 of chicken basic FGF receptor. Positive clones were inserted
CC into pUC18/119 and the recombinant plasmids were used to transform
CC E.coli MW1184. Single-stranded DNA was prepared from two of the
CC transformants for use as a template in site-directed mutagenesis.
CC Eventually the sequence corresponding to the extracellular domain
CC was cloned into plasmid pET3c under the control of the T7 promoter.
CC This amino acid sequence was deduced from the nucleotide sequence
CC of the insert from plasmid pTB1390. The transformant E.coli
CC MM294(DE3)/pLys5S, pTB1290 (FERM BP-3217) is claimed. See also AAQ14845
XX
SQ Sequence 355 AA:
XX
XX
alignment_scores:
XX Quality: 437.50 Length: 338
XX Ratio: 2.093 Gaps: 9
XX Percent Similarity: 61.834 Percent Identity: 32.544
XX
XX
alignment_block:
XX US-09-598-042A-2 x AAR15264 ..
XX
XX
Align seg 1/1 to: AAR15264 from: 1 to: 355
XX
100 GTGGCCGGCTGGGGCCGCACTATGCGGCTGCAGTCCGAGTGAGGGGGA 149
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 31 ValAlaAlaProGlyGlnSerLeuGluValArgCysLeuLeuS.... 45
XX : : : : : : : : : : : : : : : : : : : : : : : : :
XX 150 CCGCGCGCGCTGCACATGTGACCAAGATGGCCGACCATTCACACAG 199
XX : : : : : : : : : : : : : : : : : : : : : : : : :
XX 46 AspAlaAlaValIleSerTrpTrpHisAspGly.....ValHisLeuG 60
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 200 GCTGGAGCGCGCTTCCGGCTGCTG...CCGAGAGGGGCTGAAGGAGAC 246
XX || : : : : : : ||||| : : : ||||| : : : |||||
XX 60 LyrProAsnAsnArgThrValLeuIleGlyIuTrpLeuGlnIleGlyS 76
XX 247 GTGGAGCGGGAGATGCCGGCGTGTACGTGTGCAAGCCACCAACAG 296
XX : : : : : : : : : : : : : : : : : : : : : : : : :
XX 77 AlaThrProArgAspSerGlyLeuTrpValCysThrAlaSerArgThr 93
XX : : : : : : : : : : : : : : : : : : : : : : : : :
XX 297 CGGACGCTGAGCGTCACTACACCTCGTCGTCGATGACATTAGC 346
XX || : : : : : : : : : : : : : : : : : : : : : : :
XX 93 LAspSerGluThrTrpTrpArgPheMetValAsnValThrAspAlaIle 110
XX 347 CAGGAGAGGAGAGCGTGGGGCCGACAGAGCGCCCTGGGGTCAAGAGAC 396
XX || : : : : : : : : : : : : : : : : : : : : : : :
XX 110 erGlyAspAsp.....GluSpAspThrAspAlaGluAsp 122
XX : : : : : : : : : : : : : : : : : : : : : : : : :
XX 397 CCGCGCACGCCAG.....CAGTGGGACGACCGCGCTTACACAGCC 437
XX : : : : : : : : : : : : : : : : : : : : : : : : :
XX 123 PheValSerGluAsnSerAsnAsnValArgAlaProTyrTrpThrAsn 139
XX 438 CTCCAAGATGAGCGCGCGGTGATCGACGCGCCCGTGGTACTCGCTGC 487
XX : : : : : : : : : : : : : : : : : : : : : : : : :
XX 139 rGluValSerGluValArgPheHisAlaValProAlaAlaAsnTrpVal 156

```

FT	Domain	1..374	/label= extracellular domain
FT	Domain	375..395	/label= transmembrane domain
FT	Domain	396..819	/label= intracellular domain
FT	Domain	54..100	/label= Igi
FT	Disulfide-bond	54..100	
FT	Region	65	/label= conserved Trp
FT	Region	94..100	/label= consensus sequence
FT	Region	125..132	/note= "DXGXXC"
FT	Region	125..132	/label= acidic segment
FT	Domain	176..228	/note= "unusual"
FT	Disulfide-bond	176..228	
FT	Region	188	/label= conserved Trp
FT	Region	222..228	/label= consensus sequence
FT	Region	275..339	/note= "DXGXXC"
FT	Domain	275..339	/label= IgiII
FT	Disulfide-bond	275..339	
FT	Region	287	/label= conserved Trp
FT	Region	333..339	/label= consensus sequence
FT	Region	396..482	/note= "DXGXXC"
FT	Region	396..482	/label= juxtamembrane region
FT	Region	482..759	/note= "unusually long"
FT	Region	482..759	/label= TK region
FT	Binding-site	483..488	/note= "split by 14 AA insertion"
FT	Binding-site	483..488	/label= ATP binding site
FT	Binding-site	512	/note= "consensus: GXGXG"
FT	Region	619..628	/label= conserved Lys
FT	Region	639..644	/label= TK motif
FT	Region	639..644	/label= TK motif
FT	Region	651	/label= conserved Tyr
FT	Region	651	/note= "analogous to major phosphoryln. site of pp60(v-src)[Tyr 416]"
FT	Peptide	35..53	/label= tryptic peptide
FT	Peptide	56..67	/label= tryptic peptide
FT	Peptide	139..158	/label= tryptic peptide
FT	Peptide	139..158	/label= tryptic peptide
FT	Peptide	208..223	/label= tryptic peptide
FT	Peptide	264..276	/label= tryptic peptide
FT	Peptide	264..276	/label= tryptic peptide
FT	Peptide	277..289	/note= "used to design probe"
FT	Peptide	298..310	/label= tryptic peptide
FT	Peptide	298..310	/note= "used to design probe"
FT	Peptide	404..414	/label= tryptic peptide
FT	Peptide	476..493	/label= tryptic peptide



FT	/label= tryptic peptide
FT	/note="homologous to consensus for TK domain"
FT	588..596
FT	Peptide
FT	/label= tryptic peptide
FT	637..644
FT	Peptide
FT	/label= tryptic peptide
FT	/note="homologous to consensus for TK domain"
FT	665..673
FT	Peptide
FT	/label= tryptic peptide
FT	/note="homologous to consensus for TK domain"
FT	674..686
FT	Peptide
FT	/label= tryptic peptide
FT	/note="homologous to consensus for TK domain"
FT	755..776
FT	Peptide
FT	/label= tryptic peptide
FT	76
FT	Modified-site
FT	/label= N-glycos_site
FT	116
FT	Modified-site
FT	/label= N-glycos_site
FT	183
FT	Modified-site
FT	/label= N-glycos_site
FT	224
FT	Modified-site
FT	/label= N-glycos_site
FT	238
FT	Modified-site
FT	/label= N-glycos_site
FT	262
FT	Modified-site
FT	/label= N-glycos_site
FT	294
FT	Modified-site
FT	/label= N-glycos_site
FT	315
FT	Modified-site
FT	/label= N-glycos_site
FT	328
FT	Modified-site
FT	/label= N-glycos_site
FT	584
FT	Modified-site
FT	/label= N-glycos_site
FT	722
FT	Modified-site
FT	/label= N-glycos_site
PN	WO9100916-A.
PD	24-JAN-1991.
XX	
XX	06-JUL-1990;
XX	90WO-US03830.
XX	06-JUL-1989;
XX	89US-0377003.
XX	
XX	(REGC ) UNIV OF CALIFORNIA.
XX	
XX	Williams LT, Johnson DE, Lee PE;
XX	
XX	WPI: 1991-051340/07.
XX	P-PSDB: AAR10648.
XX	
XX	Fibroblast growth factor receptors - and clones encoding them,
XX	sol. forms of receptors and antibodies to them useful for
XX	diagnostic studies and therapy.
XX	
XX	Disclosure: Fig 3; 8app; English.
XX	
XX	The sequence was deduced from a cDNA clone isolated from a library
XX	prep'd. from mRNA from chicken embryos (day 6). It encodes a basic
XX	fibroblast growth factor receptor which can be used for tissue
XX	repair or development and for neuron maintenance. The DNA and
XX	protein and antibodies prep'd. from it can be used to treat
XX	diseases resulting from aberrant FGF expression and for elucidating
XX	the mechanism of FGF and identifying agonists for the prevention of
XX	angiogenesis and control of tumours.
XX	See also AAR10648.
XX	
XX	Sequence 819 AA:
XX	

[illegible]









```

462 LuleuprogliuAspproArg 468
seq_name: /STDSL/gcgdata/geneseq/geneseq/AA1994.DAT:AAV21643
seq_documentation_block:
ID   AAV21643 standard; Protein; 351 AA.
XX
AC   AAV21643:
XX
DT   16-AUG-1999 (first entry)
XX
DE   FGFR 1 (flg 5) extracellular domain.
XX
KW   Neoplastic disease; fibroblast growth factor receptor; FGFR, human;
    malignant melanoma; extracellular domain; flg 5.
XX
OS   Homo sapiens.
XX
PN   WO9400599-A2.
XX
PD   06-JAN-1994.
XX
PE   14-JUN-1993; 93MO-US05703.
XX
PR   18-JUN-1992; 92US-090646.
XX
PR   02-JUN-1995; 95US-0459296.
XX
PA   (PRIZ-) PRIZM PHARM INC.
PA   (WHIT-) WHITTIER INST DIABETES & ENDOCRINOLOGY.
XX
PI   Baird JA, Gonzalez A, Nova MP.
XX
WP: 1994-026234/03.
XX
PT   Detecting neoplastic disease - by detecting prod. indicative of
PT   elevated amts. of fibroblast growth factor receptor of gene
PT   expression
XX
PS   Claim 26; Page 23-24; 30pp; English.
XX
CC   The invention relates to a process for detecting neoplastic disease in a
CC   subject. The method comprises detecting in a sample obtained from the
CC   subject a product indicative of elevated expression of a fibroblast
CC   growth factor receptor (FGFR) gene or its product, where detection of the
CC   product indicates the presence of neoplastic disease in the subject. The
CC   method can be used for the early detection of neoplastic disease,
CC   particularly malignant melanoma. The present sequence represents the
CC   extracellular domain of FGFR 1 (also termed as flg 5).
XX
SO   Sequence 351 AA;

alignment_scores:
    Quality: 425.50      Length: 321
    Ratio: 1.970         Gaps: 9
    Percent Similarity: 67.290      Percent Identity: 31.776

alignment_block:
US-09-598-042A-2 x AAV21643 ..
.

Align seg 1/1 to: AAV21643 from: 1 to: 351

112 GCGCGCATGATGCGGTGAGTGCAGTGCAGTGGAGGGAGCCGCGCGCGT 161
||| :|||:|||||:|||| :||| :|||
25 GYAspLeuIeuGlnIeuArgCysArgIeuArgAspValGlnSerI 41
162 GACCATGTGACCAAGATGCGCGCAGCCATCCACAGCGCGTGAAGCCGCT 211
::: ||| :|||:||||| :||| :|||
41 easn...TprIeuArgAspGlyValGlnIeuAlaGlnSer...AsnArgT 56

212 TCCGCGTCTGCTGCCGCGAGGGCTGAAGGTGAAGCAAGTGCAGCGGAGAGT 261
||||: :||| :|||:|||||:||||| :|||
56 hrrArgIleThrGlyGlnGluValGluValGlnAspSerValProAlaasp 72

```

```

262 GCCGGCGGTGTACGTGTGCAGGCCCAACCAAGGGCTTCGGAGCGCTGAGCG 311
      :::::::::::::::::::: .....
73  SerGlyLeuTyrAlaCysValThrSerSerProSerGlySer...AspIn 88
      ::::|::|  ||  ::::::::::|  ::::::::::::::|::|::|::|::|
312  CAACATACACCCCTCGTCGTGTGATGACATCATGACCCAGGAGGAGGAGGCC 361
      ::::|::|  ||  ::::::::::|  ::::::::::::::|::|::|::|::|
88  rThrTyrThreserValAsnValSerAspAlaLeuProSerSerGluAla 105
      ::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|
362  TGGGGCCGACAGCTCTCTGGGGGTCAAGAG.....GACCCCGC 402
      ::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|
105  sPAspAspAspAspAspSerSerSerSerGluGluLysGluThrAspAsnThr 121
      ::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|
403  AGCCAGCAGTGGGCAAGCAAGCGGGCTTCACACAGACCCCTCAAGATGAGCG 452
      ::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|
122  LysProAsnProValAlaProTyrTrpThrSerProGluLysMetGlu 138
      ::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|
453  CCGGGTGTATCGCAGCGCCCGGTGGTAGCTCCGTGCGGCTCAAGTGGCTGG 502
      ::::::::::|  ||  ::::::::::|  ||  ::::::::::|  ||  ::
138  sLysLeuHisAlaValProAlaAlaLysThrValLysPheLysCysPro 155
      ::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|
503  CCAGCGGGCACCCTGGCGCCGACATCATCGTGTGATGAGGAGGAGCAGCGCC 552
      ::::::::::|  ||  ::::::::::|  ||  ::::::::::|  ||  ::
155  eSerGlyThrProAsnProThrLeuAlaGlyTrpLeuLysAsnGlyLysGlu 171
      ::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|
553  TTGAGCGCGCCGAGAG.....GCCGTGAGCCCGCAGGAAGAA 590
      ::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|
172  Phe.....LysProAspHisArgIleGlyGlyTyrLysValArgIleAlaThr 187
      ::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|
591  GTGACACTGAGCTGTGAGAACCTCGCGCGCGAGGACAGCGGCAATATACA 640
      ::::::::::|  ||  ::::::::::|  ||  ::::::::::|  ||  ::
187  rTrpSerIleIleMetAspSerValValProSerAspLysGlyAsnTyr 204
      ::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|
641  CCTGCCCGCTGTGCAGACCGCGCGGGCGCATCAAGCCACTTCACAAAGTG 690
      ::::|::|  ||::|::|::|::|::|::|::|::|::|::|::|::|
204  hCysLLeValGluAsnGluTyrGlySerLLeAsnHisThrTyrGluLeu 220
      ::::|::|  ||::|::|::|::|::|::|::|::|::|::|::|::|
691  GATGTGATGCAGCGAGCCCGGTTCACAGCCCGTGTCTACAGGCAAGCAACC 740
      ::::|::|  ||::|::|::|::|::|::|::|::|::|::|::|::|
221  AspValValGluArgSerProHisArgProIleLeuGlnAlaGlyLeuPr 237
      ::::|::|  ||::|::|::|::|::|::|::|::|::|::|::|::|
741  CGTGAACACAGCGGTGACTCTGGGGGAGACACAGTCTCTTCAGTGCAGG 790
      ::::|::|  ||::|::|::|::|::|::|::|::|::|::|::|::|
237  oAlaAsnLysThrValAlaLeuGlySerAsnValGluPheMetCysLysV 254
      ::::|::|  ||::|::|::|::|::|::|::|::|::|::|::|::|
791  TGGCGAGGACAGGAGACCGGTATGCAGAGGCTGAGGAGCCGTGTGAGTAC 840
      ::::|::|  ||::|::|::|::|::|::|::|::|::|::|::|::|
254  aTyrSerAspProGlnProHisIleGlnTrpLeuLysHisIleGluVal 270
      ::::|::|  ||::|::|::|::|::|::|::|::|::|::|::|::|
841  GGGCGCCGAGGGCCGACCAATCCACCATGATGTGGCGGGCCAGAAATT 890
      ::::|::|  ||::|::|::|::|::|::|::|::|::|::|::|::|
271  .....AsnGlySerLysIleGlyProAspAsnLeuProTyr 282
      ::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|
891  TGTG...GTGTCCCAACGGGTATCGTGTGTGGCGGCCGAGCGGTCT 937
      ::::|::|  ||::|::|::|::|::|::|::|::|::|::|::|::|
282  rValGlnIleLeuLysThrAlaGlyAla.....AsaThrThrAspLysG 297
      ::::|::|  ||::|::|::|::|::|::|::|::|::|::|::|::|
938  ACGTCATATAGCTGCATACACCCGTGCCCGCAGAGAGATGCGGGCATG 987
      ::::|::|  ||::|::|::|::|::|::|::|::|::|::|::|::|
297  LuMetGluValLeuHisLysLeuArgAsnValSerPheGluAspAlaGlyGlu 313
      ::::|::|  ||::|::|::|::|::|::|::|::|::|::|::|::|
988  TACATGTGCTTGGCGCAACACCATGGGTACAGTTCGCGAGCGCTT 1037
      ::::|::|  ||::|::|::|::|::|::|::|::|::|::|::|::|
314  TyrThrCysLeuAlaGlyAsnSerIleGlyLeuSerHisHisSerAlaTr 330
      ::::|::|  ||::|::|::|::|::|::|::|::|::|::|::|::|
1038  CCGTACCCGTGCTG 1050
      ::::|::|  ||::|::|::|::|::|::|::|::|::|::|::|::|
330  pleuThrValLeu 334
      ::::|::|  ||::|::|::|::|::|::|::|::|::|::|::|::|
seq_name: /SIDS1/gcdata/geneseq/geneseq/A1191.DAT.A1R13549
seq_documentation_block:
ID  A1R13549 standard: Protein: 822 AA.

```

```

XX AAR13549;
AC
XX
XX 28-OCT-1991 (first entry)
DT
XX Basic FGF receptor.
DE
XX
XX Basic fibroblast growth factor; human.
KM
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Peptide 1..21
FT Protein /label= signal_sequence
FT /label= mat_peptide
FT Region /label= extracellular_portion
XX
XX W09111459-A.
PN
XX 08-AUG-1991.
PD
XX 21-JAN-1991; 91WO-EP00103.
PE
XX 23-JAN-1990; 90GB-0001466.
PR
XX (FARM ) FARMITALIA C ERBA SRL.
PA
XX
XX Bergonzoni L, Mazue G, Isacchi A, Roncucci R, Sarmientos P;
PI WPI: 1991-252611/34.
DR N-PSDB; AAQ13311.
XX
XX Extracellular form of human fibroblast growth factor receptor
PT used to treat tumours, abnormal angiogenesis e.g. diabetic
PT retinopathy, rheumatoid arthritis and arteriosclerosis and as
PT contraceptives.
XX
XX Claim 1; Fig 3; 29pp; English.
PS
XX
XX The sequence was deduced from the cDNA sequence obid. from two
XX overlapping clones, p15 and p10, isolated from a placental lambda
XX g111 cDNA library. The DNA can be used to express recombinant
XX bFGF receptor which is an antagonist of human aFGF and bFGF. The
XX receptor can be used to treat abnormal angiogenesis (e.g. in
XX diabetic retinopathy, neovascular glaucoma, etc. and possibly
XX certain solid tumours), and in contraceptives. Dosage is 10-100 ug.
XX
SQ Sequence 822 AA;

Alignment_scores:
      Quality: 423.50      Length: 442
      Ratio: 1.580      Gaps: 14
      Percent Similarity: 60.633      Percent Identity: 27.602

Alignment_block:
US-09-598-042A-2 x AAR13549 ..
Align seg 1/1 to: AAR13549 from: 1 to: 822

112 GCGCCACATATGCGGTGACAGTGCACATGAGAGGGGACCCGCCCGCT 161
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
48 GYASPALEuLEuGlnLEuAArgYsARgLEuARgAsPAsPValGlnSer11 64
162 GACCATGTGGACCAAGATGCGCCGACCATCCACAGCGGCTGGAGCCGCT 211
: : : : : : : : : : : : : : : : : : : : : : : : : :
64 eAsN...tRLEuARgAsPAsPValGlnLEuAAGLusER...AsnARgT 79
212 TCCGGGTGTGCGCGAGGGGCTGAAGGTGAAGCAAGGTGAGGGAGAGAT 261
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
79 hAArgLEuHrGLuGlnGluValGlnValGlnAsPSeRValProAlAsP 95

```

```

262 GCGGGGTGTAGTGTGACAGGCCAACACAGGCTTGGCAGCGCTAGAGCT 311
: : : : : : : : : : : : : : : : : : : : : : : : : :
96 SerGlyLEuTYrAlaCysValThrSerSerProSerGlySer...AspH 111
312 CAACATACACCCCTGCTGCTGATGACATGTAGCCAGGAGAGAGAGAGC 361
: : : : : : : : : : : : : : : : : : : : : : : : : :
111 tThrTYrPhSerValAsnValSerAsPAsPAlaLEuProSerSerGluAsP 128
362 TGGGGCCCGACAGCTCTCTGGGGGTCAAGAGACCCCGCC..... 402
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
128 sPAsPAsPAsPAsPAsPAsPAsPAsPAsPAsPAsPAsPAsPAsPAsP 144
403 .....AGCCACAGTGGGACAGACCGCGCTTCACACAGACCCCTCAAGAT 446
: : : : : : : : : : : : : : : : : : : : : : : : : :
145 LysProAsnArgMeCProValAlaProTYrTYrTYrSerProGluLysMe 161
447 GAGGGCGCGGTGATGACGACGCGCGGTGGTATGCTCCGGCGGCTCAAGT 496
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
161 tGluLysLysLEuHISAlaValProAlaLysThrValLysPheLysC 178
497 GCGTGGCCAGCGGACACCTCGCCGACATCAGCGTGAAGAGAGACGAC 546
|| : : : : : : : : : : : : : : : : : : : : : : : : : :
178 YsProSerSerGlyThrProAsnProThrLEuArgTYrLEuLysAsnGly 194
547 CAGGCTTGACGCGCCACAG.....GCCGCTGAGCCACAGAA 584
: : : : : : : : : : : : : : : : : : : : : : : : : :
195 LysGluPhe...LysProAsPAsArgILEgLYGlyTYrLysValArgTY 210
585 GAAGAAGTGAACATGAGCGCTGACAGACCTGGCGGAGGACAGACGGA 634
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
210 rAlaThrTYrSerILElleMetAsPSeRValAlaProSerAsPAsGlyA 227
635 AATACACCTGCGCGCTGTGACACCGCGGCGCCATCAACGCCACCTAC 684
: : : : : : : : : : : : : : : : : : : : : : : : : :
227 sNTYrTYrCysILEValGluAsnGluTYrGlySerILEsntHISntHyr 243
685 AAGGTGATGTATCCAGCGGACCCGTTCCAAAGCCGTTGCTCACAGGAC 734
: : : : : : : : : : : : : : : : : : : : : : : : : :
244 GlnLEuAsPAsValAlaGluArgSerProHISArgProILEuGlnAla 260
735 GCACCCCGGACACAGAGCGTGTGACTTGGGGGACACAGTCTCTCCAGT 784
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
260 YLEuProAlaAsnLysThrValAlaLEuLysSerAsnValGluPheMeC 277
785 GGAAGTGTGGCGAGCAGCGTGAAGCGGTGATCAGTGGTGAAGCGGCG 834
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
277 YsLysValTYrSerAsPProGlnProHISILEgINTPrLEuLysHISILE 293
835 GAGTACGGCGCGAGAGCGCCACAACTCCACATCGATGTGGCGGCA 884
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
294 GluVal.....AsnGlySerLysILEgLYProAsPAsPhe 305
885 GAACTTTGTG...GTGCTGCCAGGAGTGAAGTGTGTCGCGCGCGAG 931
: : : : : : : : : : : : : : : : : : : : : : : : : :
305 uProTYrValGlnILEuLysTrLEuGlyVal.....AsnThrThra 320
932 GCTCTACTCAATAAGCTGCTCATACCCGCGCCGCGGACAGATGG 981
: : : : : : : : : : : : : : : : : : : : : : : : : :
320 sPLYsGluMeTGLuValLEuHISLEuAArgAsnValSerPheGluAsPAla 336
982 GGCATGTACATCTGCTTGGCGCCCAACACACATGAGGCTACGCTCGGAG 1031
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
337 GLyGluTYrTYrCysLEuAlaGLYsntSerILEgLYLEuSerHISISe 353
1032 CGCCTTCTCACCGCTGCTGCCA.....GACCCAAACCGCGAGGCGCAC 1075
||| : : : : : : : : : : : : : : : : : : : : : : : : : :
353 rAlaTrPrLEuThrValLEuGluAlaLEuGlnGluTYrPro..... 366
1076 CTGTGGCCTCGTCGTCGCGCACTAGCGTGGCGCGCGGTGCTATC 1125
367 .....AlaValMeThrSerProLEuTYrLEuGlnLEuILE 378

```





```

210 rAlaThrTrpSerIleIleMetAspSerValAlaProSerAspLysGlyA 227
XX
635 AATACACCTGCGCGCTGTGGAACCGCGCGGCCCATCAACGCCACTAC 684
DE
227 sntYrThrcysIleValGluAsnGluTrpGlySerIleAsnHsTrpTrp 243
XX
685 AAGGTGATGTATCCAGCGGACCGCTCCAGCCCGCTGCTCAGACGAC 734
OS
244 GluLeuAspValValGluArgSerProHsArgProIleLeuGlnAla 260
XX
735 GCACCCCGTGAACACGACGCTGACTCGGGGAGACACAGCTCTTCAGT 784
PD
260 yLeuProAlaAsnLysThrValAlaLeuGlySerAsnValGluPheMetC 277
XX
785 GCAAGGTGCGCAGCGACGTGAAGCCGGTATCCAGTGGCTGAAGCCGG 834
XX
277 yLysValLysTrpSerAspProGlnProHsIleGlnTrpLeuHsIle 293
XX
835 GAGTACGGCGCGGAGCGCGCCCAACACTCCACCATGATGTGGCGGCA 884
DR
294 GluVal.....AsnGlySerLysIleGlyProAspAsnLe 305
DR
885 GAAGTTTGTG...GTGCTGCCACGCGGTGACGTGTGGTGGCGCGGAG 931
PT
305 uProTrpValGlnIleLeuLysThrAlaGlyAla.....AsnThrThra 320
XX
932 GCTCTACCTCAATAAGCTGTCTCATCACCCTGCGCGCCGACGACGATG 981
XX
320 sPluGlnIleuGluValLeuHsIleuAlaArgAsnValSerPheGluAsp 336
XX
982 GGCATGTATCATGCTGGCGCCCAACACCATGGCTACAGCTCCGCGAG 1031
XX
337 GlyGluTrpThrCysLeuAlaGlyAsnSerIleGlyLeuSerHsIleS 353
XX
1032 CCGCTCTCTACCGCTGTGCTGCA.....GACCCAAACCGCCAGCGC 1075
SQ
353 rAlaTrpLeuThrValLeuGlnAlaLeuGluAlaArgPro..... 366
XX
1076 CTGTGGCTCTCTGCTCGGCGCACTAGCTGCGCTGGCGCGCTGTATC 1125
XX
367 .....AlaValMetThrSerProLeuTrpLeuGluIleIle 378
XX
1126 GGCATCCCGAGCGCGCGCTGTCTATC.....CTGGGACCGCTGT 1166
XX
379 IleTrpCysThrGlyAlaPheLeuIleSerCysMetValGlySerVal 395
XX
1167 CCTGTGGCTTTGCCAGGCCCAAGAACAGCG..... 1197
XX
395 eValTrpLysMetLysSerClyThrLysLysSerAspPheHsIleSerG 412
XX
1198 ..TGCACCCCGCGCGCTGCTCCCTGCTGGGCGACCGCGCGCGGG 1245
XX
412 eValValHsIleLysLeuAlaLysSerIleProLeuArgArgGlnVal 428
XX
1246 AGGGCCCGGACCGGACGCGGACACAAAGAC.....CT 1277
XX
429 ValSerAlaAspSerSerAlaSerMetAsnSerGlyValLeuLeuVal 445
XX
1278 TCCCTGCTGTGGCGCGCTGAGCGCTGCGCTGCTGTGGGGCTGTGAGG 1327
XX
445 gProSerArgLeuSerSerSerClyThrProMetLeuAlaGlyValSer 462
XX
1328 AGCATGGGTCTCCGCGACGCCCCCAAG 1353
XX
462 LuTrpGluLeuProGluAspProArg 470
XX
seq_name: /SISL/gcdata/geneseq/geneseqp/AA1992.DAT:AA26337
seq_documentation_block:
ID AAR26337 standard; Protein; 822 AA.
XX
AC AAR26337;
XX

```

```

DT 03-FEB-1993 (first entry)
XX
DE N-sam.
XX
KW Tumour gene; bFGF receptor; antibody; detection; drug.
XX
OS Homo sapiens.
XX
PN JP04190792-A.
XX
PD 09-JUL-1992.
XX
PE 22-NOV-1990; 90JP-0316100.
XX
PR 22-NOV-1990; 90JP-0316100.
XX
PS (KOKU-) KOKURITSU GAN CENT SOCHO.
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
DR WPI, 1992-280111/34.
DR N-PSDB; AA027658.
XX
PT Gene prod of N-sam tumour - and its recombinant-contg. nucleotide
sequence
XX
PS Claim 1; Fig 1; 19pp; Japanese.
XX
CC The sequence given encoded by the N-sam tumour gene. This protein
CC could be the receptor of human bFGF. The N-sam protein and peptide
CC fragments of N-sam are useful as anti-tumour drugs. Anti-N-sam peptide
CC antibodies are useful for the detection of the N-sam gene product, in
CC the purification of the N-sam gene product and in the diagnosis of
XX tumours and anti-tumour drugs.
XX
SQ Sequence 822 AA;
XX
alignment_scores:
Ratio: 423.50 Length: 442
Percent Similarity: 60.633 Percent Identity: 27.602
XX
alignment_block:
US-09-598-042a-2 x AAR26337 ..
XX
Align seg 1/1 to: AAR26337 from: 1 to: 822.
XX
112 GGCAGCATATGCGGCTGCGAGTGCACGAGGAGGGGACCGCGCGCGCT 161
48 GlyAspLeuLeuGlnLeuArgCysArgLeuArgAspAspValGlnSerIle 64
162 GACCATGTGAGACCAAGATGCGCGCACCATTCACACAGCGCTGAGCGCGCT 211
64 eAsn...TrpLeuArgAspGlyValGlnLeuAlaGluSer...AsnArgTr 79
212 TCCGCTGCTGCCCGCAGGGCTGAAGTGAAGCAGTGAAGCGCGGAGAT 261
79 hrArgIleThrGlyGlnGluValGlnValGlnAspSerValProAlaAsp 95
262 GCGCGGCTGAGTGTGCAAGCGCACCAACGCGCTGCGAGCGCTGAGCGT 311
96 SerGlyLeuTrpAlaCysValThrSerProSerGlySer...AspH 111
312 CAATACACCTCTGCTGCTGATGATTCAGCCAGGAGAGAGAGAGCC 361
111 rThrTrpPheSerValAsnValSerAspAlaLeuProSerSerGluAspA 128
362 TGGGCGCGCAGACTCTCTGGGGGTCAAGAGAGACCGCGC..... 402
128 sPAspAspAspAspAspSerSerSerGlnGluLysGluThrAspAsnThr 144
403 .....AGCCAGAGTGGGACGACGCGCGCTTCACACAGCGCTCAAGAT 446

```

```

145  LysProAsnArgMetProValAlaProTyrThrSerProGluLysMe 161
447  GAGCGCGCGGTGATCGACGCGCGGTGGAGTCCGTGGGCTCAAGT 496
161  TGIuLysLysLeuHisAlaValProAlaAlaLysThrValLysPheLysC 178
497  GCGTGGCCAGCGACCCCTCGCGCCGACATCATGATGATGATGAGACGAC 546
178  ySProSerSerGlyThrProAsnProThrLeuArgTyrPheLysAsnGly 194
547  CAGGCTTGCAGCGCCAGAG.....GCCGCTGAGCCCGAGAA 584
195  LysGluPhe...LysProAspHisArgIleGlyLysValArgTyr 210
585  GAAGAAGTGCACAGTACGACCTGAGAACCTGCGCGGAGGACGCGGCA 634
210  rAlaThrThrSerIleIleMetAspSerValValProSerLysGly 227
635  AATACACCTCGCGCGTGTGACACCGCGCGGACCATCAACGCGCATAC 684
227  snTyrThrCysIleValGlnAsnGlyTyrGlySerIleAsnHisThrTyr 243
685  AAGGTGATGTGATTCACAGCGGACCGGTTCACAGCCGTGTCACAGGAC 734
244  GlnLeuAspValValGluArgSerProHisArgProIleLeuGlnAlaG 260
735  GCACCCGCGTGAACAGACGCGTGAAGTTCGCGGAGACGACGCGCTTCA 784
260  yLeuProAlaAsnLysThrValAlaLeuGlySerAsnValGluPheMetC 277
785  GCAAGTTCGCGACGCGACGTGAAGCCGCGTATCATGCGTGAAGCGCTG 834
277  yLysValTyrSerAspProGlnProHisIleGlnThrPheLysHisIle 293
835  GAGTACGCGCGCGAGGCGCGCACACTCCACATCATGATGCGCGGCGCA 884
294  GluVal.....AsnGlySerLysIleGlyProAspAsnIle 305
885  GAAGTTGTG...GTGCTGCCACGCGGTGAGTGTGCGGCGCGCGACG 931
305  uProTyrValGlnIleLeuLysThrAlaGlyVal.....AsnThrThr 320
932  GCTCTTACCTCAATAGCTGCTCATACCGCGTCCCGCGGACGAGATGCG 981
320  sPlysGlnMetGluValLeuHisLeuArgAsnValSerPheGlnAspAla 336
982  GGCATGTACATCTGCTTGGCGGACACCATGGGCTACAGCTTCCGACG 1031
337  GlyIuTyrThrCysLeuAlaGlyAsnSerIleGlyLeuSerHisHisSe 353
1032  CGCCTTCTCACCCTGTGCCA.....GACCCAAACCGCGCAGGCGCAC 1075
353  rAlaThrPheThrValLeuGlnAlaLeuGlnGluArgPro..... 366
1076  CTGTGGCCTCTCGTCCGCGGACCTAGCCTGCGTGGCCCGGTGTCATC 1125
367  .....AlaValMetThrSerProLeuTyrLeuGlnIleLe 378
1126  GGCATCCACGCGCGCTGTCTTCATC.....CTGGGACCCCTCTCT 1166
379  IleTyrCysThrGlyAlaPheLeuIleSerCysMetValGlySerValI 395
1167  COTGGGGCTTGGCCGCGGACGAGAACGCG..... 1197
395  eValIuTyrLysMetLysSerGlyThrLysLysSerAspPheHisSerGln 412
1198  ..TGCACCCCGCGCGCTCCCTCCCTGCTGGGACGCGCGCGCGGGG 1245
412  eValAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 428
1246  ACGGCGCGCGACGCGGAGACAGAC.....CT 1277
429  ValSerAlaAspSerSerAlaSerMetAsnSerGlyValLeuLeuValAr 445

```

```

1278  TCCTCGTGGCGCGCCTCAAGCGCTGCGCGTGGTGGCGTGTGAGG 1327
445  gProSerArgLeuSerSerSerGlyThrProMetLeuAlaGlyValSerG 462
1328  AGCATGGGTCTCCGCGAGCGCCCGACG 1353
462  LuTyrGluLeuProGluAspProArg 470

seq_name: /STDS1/gcgdata/geneseq/geneseq/AA1998.DAT.AAM63844
seq_documentation_block:
ID   AAM63844 standard; Protein; 820 AA.
AC   AAM63844;
DT   01-OCT-1998 (first entry)
DE   Mouse bFGF receptor protein.
KW   Basic fibroblast growth factor receptor; bFGF; heparin binding; murine;
    antitumour agent; inhibitor; wound healing.
OS   Mus sp.
    Key      Location/Qualifiers
    FH      Misc-difference 67
    FT      /label= unknown
    FT      Misc-difference 68
    FT      /label= unknown
    FT      Misc-difference 260
    FT      /label= unknown
    FT      Misc-difference 281
    FT      /label= unknown
    FT      Misc-difference 282
    FT      /label= unknown
    FT      Misc-difference 576
    FT      /label= unknown
    FT      Misc-difference 702
    FT      /label= unknown
    FT      Misc-difference 702
    FT      /label= unknown
    PN      US579182-A.
    PD      04-AUG-1998.
    PF      14-DEC-1993. 90US-0166717.
    PR      20-DEC-1990. 90US-0631717.
    PR      14-DEC-1993. 930S-0166717.
    PA      (CHIL-) CHILDRENS MEDICAL CENT.
    PA      (HARD ) HARVARD COLLEGE.
    PI      Flanagan JG, Klagsbrun M, Leder P, Ornitz DM, Yason A;
    DR      WPI; 1998-446074/38.
    DR      N-PSDB; AAV44041.
    PT      Assays for high-affinity heparin-binding growth factor receptor
    PS      ligands - using receptor-overexpressing cells or cell-free system
    PS      Disclosure; Fig 9; 38pp; English.
    XX      This sequence represents a murine basic fibroblast growth factor (bFGF)
    CC      which is a member of the heparin-binding growth factor receptor family.
    CC      This protein is used in a method which assays the ability of a substance
    CC      to bind to a high-affinity heparin-binding growth factor (bFGF) receptor.
    CC      The assay screens for potential antitumour agents that inhibit binding of
    CC      bFGF to high-affinity receptors, or for potential wound healing agents
    CC      that promote such binding.
    SQ      Sequence      820 AA:

```

## alignment\_scores:

Quality: 413.50 Length: 441  
 Ratio: 1.560 Gaps: 15  
 Percent Similarity: 60.091 Percent Identity: 27.891

## alignment\_block:

US-09-598-042a-2 x AAW63844 ..

Align seg 1/1 to: AAW63844 from: 1 to: 820

```

112 GCGCCGACATATGAGGCTGCATGCCCCAGTGGAGGGGACCCGCCCGCT 161
113 ||| : : : : : : : : : : : : : : : : : : : : : : : : :
48 GLYAspLeuLeuGlnLeuArgCysArgLeuArgAspAspValGlnSerIle 64
162 GACCATGTGGACCAAGATGCG...CGCACATCCACAGCGGCTGGAGCC 208
: : : : : : : : : : : : : : : : : : : : : : : : :
64 easn...Tyr****AspGlyValGlnLeuValGlnSer.....AsnA 78
209 GCTTCGCGCTGCTGCCGCGAGGGCTGAAGTGAAGCAGGTGGAGCGGGAG 258
||| : : : : : : : : : : : : : : : : : : : : : : : : :
78 rGThrArgIleThrGlnGlnValGlnValArgAspSerIleProAla 94
259 GATGCCGCGGTGATGCTGTCAGAGCCACACAGCGGCTTCGCGACCTGAG 308
||| : : : : : : : : : : : : : : : : : : : : : : : : :
95 AspSerGlyLeuTyrAlaCysValThrSerProSerGlySer...As 110
309 CGTCAACTACACCTTCGCTGCTGGATGACATPAGCCAGGAGAGAGA 358
: : : : : : : : : : : : : : : : : : : : : : : : :
110 pThrThrTyrPheSerValAsnValSerAspAlaLeuProSerSerGlnA 127
359 GCCTGGGGCCGACAGCTCTCTGGGGGTCAGAG.....GACCCC 399
: : : : : : : : : : : : : : : : : : : : : : : : :
127 sPAspAspAspAspAspAspSerSerSerGlnGlnLysGlnThrAspAsn 143
400 GCGACGACAGTGGGCGACGCGGCTTCACACAGCGCTCCAGATGAG 449
: : : : : : : : : : : : : : : : : : : : : : : : :
144 ThrLysProAsnProValAlaProTyrTrpThrSerProGlnLysMetI 160
450 GCGCCGCGGTATGCGACGGCCGCTGGAGTCCGCTGCGGCTCACTGCG 499
: : : : : : : : : : : : : : : : : : : : : : : : :
160 uLysLysLeuHisArgValProAlaIleValThrValLysPheLysCysP 177
500 TGGCCAGCGCGGACCTCGCGCCGACATCAGCGGATGAAGGACGACG 549
: : : : : : : : : : : : : : : : : : : : : : : : :
177 rGSerSerGlnTyrProAsnProThrLeuArgTrpLeuLysAsnGlnLys 193
550 GCCTTGACGCGCCGACAG.....GCGGCTGAGCGCCGAGAAAGA 587
: : : : : : : : : : : : : : : : : : : : : : : : :
194 GluPhe...LysProAspHisArgIleGlyLysTyrLysValArgTyrAl 209
588 GAAGTGGACACTGAGCCTGAAGAACCTGGCGCGGAGAGACAGCGCAAA 637
||| : : : : : : : : : : : : : : : : : : : : : : : : :
209 aThrTrpSerIleIleMetLaspSerValValProSerAspLysGlyAsn 226
638 ACACCTGCGCGCTGTGGAACCGCGCGGCCCATCAACGCCACTCAAG 687
||| : : : : : : : : : : : : : : : : : : : : : : : : :
226 yThrCysLleValGlnAsnGlnTyrGlySerIleAsnHisThrTyrGln 242
688 GGTGATGATGATCAGAGGACCGCTTCAGAGCCGCTGTCAGAGCGACGA 737
: : : : : : : : : : : : : : : : : : : : : : : : :
243 LeuAspValValGlnArgSerProHisArgProIleLeuGlnIleAlaL 259
738 CCGCGTGAACAGCAGCGTTCGCGGGGAGACACAGCTCTTCACATGCA 787
: : : : : : : : : : : : : : : : : : : : : : : : :
259 u**AlaAsnLysThrValAlaLeuGlnSerAsnValGlnPheMetCysL 276
788 AGGTGGCGAGCGAGCGGATGATCCAGTGGCTGAAGCGCGTGGAG 837
||| : : : : : : : : : : : : : : : : : : : : : : : : :
276 yValTyrSerAsp****ProHisIleGlnTrpLeuLysHisIleGln 292
838 TACGGCGCGGAGGCGCCACAATCCACATCGATGTGGGCGGCGAGAA 887
: : : : : : : : : : : : : : : : : : : : : : : : :

```

```

293 Val.....AsnGlySerLysIleGlyProAspAsnLeuPr 304
888 GTTGTG...GTGTCGCCAGGGTGACGTGTGGTGGCGCCGACGCT 934
: : : : : : : : : : : : : : : : : : : : : : : : :
304 oTyrValGlnIleLeuLysThrValGlyVal.....AsnThrThrAspL 319
935 CCGACCTCAATAAGCTGCTCATCACCCTGGCGCCGACGACGATGGGGC 984
: : : : : : : : : : : : : : : : : : : : : : : : :
319 yGlnMetGlnValLeuHisLeuArgAsnValSerPheGlnAspAlaGly 335
985 ANGTACATGCTGCTGGCGCCACACACATAGGCTTACAGCTTCGCGAGCC 1034
||| : : : : : : : : : : : : : : : : : : : : : : : : :
336 GluTyrThrCysLeuAlaGlnAsnSerIleGlyLeuSerHisIleSerAl 352
1035 CTTCTCACCCTGCTGCCA.....GACCCAAACCGCCAGGGCCACCTG 1078
: : : : : : : : : : : : : : : : : : : : : : : : :
352 aTrpLeuThrValLeuGlnAlaLeuGlnGlnArgPro..... 364
1079 TGGCTCTCTGCTGCGCCACATGCTGCGCGCGCGCTGCTCATCGGC 1128
: : : : : : : : : : : : : : : : : : : : : : : : :
365 .....AlaValMetThrSerProLeuTyrLeuGlnIleIle 377
1129 ATCCACAGCGCGCTGCTTTCATC.....CTGGACACCTGCTCTCT 1169
: : : : : : : : : : : : : : : : : : : : : : : : :
378 TyrCysThrGlyAlaPheLeuIleSerCysMetLeuGlySerValIleI 394
1170 GTGGCTTTCGCGCGCCAGAAAGACCG.....T 1198
: : : : : : : : : : : : : : : : : : : : : : : : :
394 eTyrLysMetLysSerGlyThrLysLysSerAspPheHisSerGlnMetA 411
1199 GCACCCCGCGCTGCGCTGCTGCGCGCGCGCGCGCGCGCGGAGAG 1248
: : : : : : : : : : : : : : : : : : : : : : : : :
411 LalaHisLysLeuAlaLysSerIleProLeuArgArgGlnValThrVal 427
1249 GCGCGCGAGCGCGAGAGCAAGAGAC.....CTTCC 1280
: : : : : : : : : : : : : : : : : : : : : : : : :
428 SerAlaAspSerSerIleAspMetAsnSerGlyValLeuLeuValArgP 444
1281 CTCGTTGGCGCGCTGACGCTGGCGCTGCTGCTGCTGCTGCTGCTGCT 1330
||| : : : : : : : : : : : : : : : : : : : : : : : : :
444 oSerArgLysSerSerSerGlyThrProMetLeuAlaGlyValSerGlnT 461
1331 ATGGGCTTCGCGAGCGCCCGCCAG 1353
: : : : : : : : : : : : : : : : : : : : : : : : :
461 yTrpLeuProGlnAspProArg 468

seq_name: /SIDSL/gcgdata/geneseq/geneseqp/AA2000.DAT: AAB58376
seq_documentation_block:
ID AAB58376 standard; Protein: 764 AA.
XX
XX AAB58376;
XX
XX
XX 14-MAR-2001 (first entry)
XX
XX Lung cancer associated polypeptide sequence SEQ ID 714.
XX
XX DE
XX
XX Human; lung cancer associated protein; neuroprotective; cytostatic;
XX cardioactive; immunomodulatory; muscular active; vulnerable;
XX gastrointestinal; nephrotropic; antineoplastic; gynecological;
XX antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
XX proliferative disorder; wound healing; infectious disease.
XX
XX Homo sapiens.
XX
XX OS
XX
XX PN WO20005180-A2.
XX
XX PD 21-SEP-2000.
XX
XX PF 08-MAR-2000; 2000WO-US05918.
XX
XX PR 12-MAR-1999; 99US-0124270.
XX
XX

```







```

239 nValSerPheGluAspAlaGlyGlyThrThrCysLeuAlaGlyAsnSerI 256
      ::::::::::::::: ||| :::::::::::::::
1013 TGGGCTACAGCTTCGCGACGCGCTTCCTACCGCTGCTGCCA.....GAC 1056
      ::||| ||| ::::::::::::::: |||
256 LeGIleuSerHisHisSerAlaTrpLeuThrValLeuGluAlaLeuGlu 272
      ::||| ::||| ::||| ::|||
1057 CCAAAACCCGACGGGCGACCTGTGGCTCTGCTCTGCGCCACTGACCT 1106
      ::||| ::||| ::||| ::|||
273 GluArgPro.....AlaValMetThrSer.ProLeuTyrL 284
      ::||| ::||| ::||| ::|||
1107 GCGGTGGCCCTGTGTCATCGCATCCGCGCGCTGTCTTCATCTGG 1156
      ::||| ::|||
284 euGIuGly.SerGIySer..... 289
1157 GCACCTCTGCTCTGTGGCTTTGGCAGCCCGAGAGAGCGGTGACCC 1206
      ::||| ::|||
290 .....ProGIyLeuGlnGlu.....ProL 296
1207 GGGCTGCCCCCTCCCTGCTGGGCGACCGCCGCGGGAGCGCCGCGCA 1256
      ::||| ::|||
296 ySerCysAspLysThrHisThrCysProProCys.....ProAla 309
1257 CCGCAGCGGAGACAAAGACCTTCCTCC 1283
310 ProGIuLeuGlnGlyGlyProSerVal 318

seq_name: /SIDS1/gcgdata/geneseq/geneseqp/AA191.DAT:AA15266
seq_documentation_block:
ID AA15266 standard; Protein: 526 AA.
XX
AC AAR15266;
XX
DT 18-FEB-1992 (first entry)
XX
DE Clone PTB1228-encoded protein with FGF receptor activity.
XX
KW Human; fibroblast growth factor; cancer.
XX
OS Homo sapiens.
XX
PN WO9117183-A.
XX
PD 14-NOV-1991.
XX
PE 25-APR-1991; 91WO-JP00557.
XX
PR 28-DEC-1990; 90JP-0415801.
PR 27-APR-1990; 90JP-0113146.
PR 31-JUL-1990; 90JP-0204438.
PR 14-SEP-1990; 90JP-0245256.
XX
PA (TAKE ) TAKEDA CHEMICAL IND KK.
XX
PI Igarashi K, Senoo M, Watanabe T;
XX
DR WPI; 1991-353723/48.
DR N-PSDB; AAQ14848.
XX
PT New mutin(s) of proteins - with fibroblast growth factor
PT receptor activity, useful for treating multiple endocrine
PT neoplasia, prostatic hypertrophy, used for diagnosis
XX
PS Example 2; Fig 3; 88pp; English.
XX
CC A cDNA library prepared from human cancer cell line Kato III mRNA
CC was screened with an oligonucleotide corresponding to amino acids
CC 529-541 of chicken basic FGF receptor. Three positive clones were
CC obtained. One was cloned into pUC118/119 to give PTB1228. The
CC sequence of PTB1228 is shorter than that of PTB1229 lacking as it
CC does nucleotides 134-478 and 1309-1314 (see AAQ14849). The amino acid
CC sequence was deduced from the insert sequence.

```

```

XX
SQ Sequence 526 AA;

alignment_scores:
      Quality: 379.50      Length: 329
      Ratio: 1.956      Gaps: 7
      Percent Similarity: 58.967      Percent Identity: 29.483

alignment_block:
us-09-598-042a-2 x AAR15266 ..
Align seg 1/1 to: AAR15266 from: 1 to: 526

265 GCGGTGTACGTGTGCAAGGCCACCAAGCGCTTCGCGAGCGCTG... 312
      ||| ::::::::::: ||| :::::::::::
5 GluArgPheIleCysLeuValValThrMetAlaThrLeuSerLeuAl 21
313 .....AACTACACCCCTGCTGCTGGATGATGATTAAGCCAGGAG 355
21 argProSerPheSerLeuVal.....GluSprThrThrLeu 28
356 AGAGCTGGGGCCCGACAGCTCTGTGGGGGTCAAGAGACCCCGCAGC 405
29 .....GluSprThrThrLeu 33
406 CAGCAGTGGGGCAGCAGCCGCGCTTCACACAGCCCTCCAGAGAGGCG 455
      ::||| ::||| ::||| ::|||
34 GluProGIuGIyAlaProTyrThrThrAsnThrGluLysMetGIuLys 50
456 GGTGATCGCAGCGCCGCTGGGTAGCTCCGTGGCGGTCAAGTGGTGCCA 505
      ::||| ::||| ::||| ::|||
50 GluHisAlaValProAlaAlaAsnThrValLysPheArgCysProAla 67
506 GCGGCGACCCCTGCGCCGACATCATCGATGATGAAGAGCAGCGGCTG 555
      ::||| ::||| ::||| ::|||
67 LysIAsnProMetProThrMetArgTrpLeuLysAsnGlyLysGluPhe 83
556 ACG.....CGCCAGAGCGCGCTGAGCCCGAGAGAGAGAGAGTGAC 596
84 LysGlnGlnHisArgIleGlyGlyTyrLysValArgAsnGlnHisTrp 100
597 ACTGAGCTGAGAGAACCTGCGCGCGGAGAGCAGCGCAATACCTGCC 646
      ::||| ::||| ::||| ::|||
100 LLeuIleMetGluSerValValProSerAspLysGlyAsnTyrThrCys 117
647 GCGTGTGAACGCGCGCGGCGCATCAAGCCACCTACAGAGGATGAG 696
      ::||| ::||| ::||| ::|||
117 AlValGluAsnGluTyrGlySerIleAsnHisThrThrHisLeuAsp 133
697 ATCCAGCGGACCCGTTCCAGCCGCTGTACAGCAGCGACGCCGTGAA 746
      ::||| ::||| ::||| ::|||
134 ValGIuArgSerProHisArgProIleLeuGlnAlaGlyLeuProAla 150
747 CAGCAGGTGAGACTTGGGGGAGCACGCTCTTCAGTGCAGAGTGCCGA 796
      ::||| ::||| ::||| ::|||
150 nAlaSerThrValValGIyGIyAspValGIuPheValLysValTyr 167
797 GCGAGGTGAACCGGTGATCCAGTGCAGTGAAGCGCTGGAG..... 837
      ::||| ::||| ::||| ::|||
167 eTrsPAlaGlnProHisIleGlnTrpIleLysHisValGIuLysAsn 183
838 .....TACGGCGCCGAGGCGCCGACCAATCCATCATGATGTGGCG 881
184 SerLysTyrGIyProAspGIy..... 190
882 CCAGAAATTGTGTGCTGCTGCCACGGGTGACGTG.....TGTGCGCG 925
191 .....LeuProTyrLeuLysValLeuLysHisSerGIyI 202
926 CCGAGCGCTCTACTCATTAAGCTGCTCATCAAGCCGTGCGCGCGAG 975
202 LeAsnSerSerAsnAlaGluValLeuAlaLeuPheAsnValThrGluAla 218

```

```

976 GATCGGGCATGTACATCTGCTTGGCCCAACACATGGGCTACAGCTT 1025
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
219 Aspalactylgltutrylleacylsvalserasntryllelglnalaas 235
1026 CCGCAGCGCCTTCTCCACCGCTGTCGACGACCCAAACCGCAGGCGCAC 1075
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
235 nglnseralatrpleuthrvalleuprolysglnlnalaProgly... 250
1076 CTGAGGCTCTCTGCTCTGCGGCCACTAGCGCTGGCGCCGCGCTGTATC 1125
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
251 .....ArglnLysglulIethrAlaserProaspTyrLeuglnlle 264
1126 GGCATCCACCGCGGCTGTCTTCATCTGCGGACCGCTCTGTGACT 1175
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
265 AlalIetYrCysIlelglyValPheleuIleAlaCysMetValValThra 281
1176 TTGCCAGGCCAGAAAGACCGCTGACCCCGCGGCT 1212
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
281 lIleLeucysArgMetLysasnThrThrLysPro 293

seq_name: /stsd1/gcgdata/geneseq/geneseq/AA191.DAT:AA15269
seq_documentation_block:
ID AA15269 standard; Protein; 652 AA.
XX
AC AA15269;
XX
DT 18-FEB-1992 (first entry)
XX
DE Clone PTB1283-encoded complete FGF receptor.
XX
KW Human; fibroblast growth factor; cancer.
XX
OS Homo sapiens.
XX
PN WO9117183-A.
XX
PD 14-NOV-1991.
XX
PF 25-APR-1991; 91MO-JP00557.
XX
PR 28-DEC-1990; 90JP-0415801.
PR 27-APR-1990; 90JP-0113146.
PR 31-JUL-1990; 90JP-0204438.
PR 14-SEP-1990; 90JP-0245256.
XX
PA (TAKE ) TAKEDA CHEMICAL IND KK.
XX
PI Igarashi K, Senoo M, Watanabe T;
XX
DR WPI; 1991-353723/48.
XX
DR N-PSDB; AAQ14851.
XX
PT New mutin(s) of proteins - with fibroblast growth factor
PT receptor activity, useful for treating multiple endocrine
PT neoplasia, prostatic hypertrophy, used for diagnosis
XX
PS Example 3; Fig 8; 88pp; English.
XX
CC A cDNA library prepared from human cancer cell line Kato III mRNA
CC was screened with an oligonucleotide corresponding to amino acids
CC 529-541 of chicken basic FGF receptor. Three positive clones were
CC obtained. One was cloned into pUC18/119 to give pTB1228 (see
CC AAQ14848). The complete FGF coding sequence was obtained by ligating
CC the insert from pTB1228 to the DNA sequence of the plasmid pTB1281
CC insert which encodes the carboxyl terminus of the FGF receptor from
CC Glu 533 onwards.
XX
SO Sequence 652 AA;

```

alignment\_scores:

```

Quality: 379.50 Length: 329
Ratio: 1.956 Gaps: 7
Percent Similarity: 58.967 Percent Identity: 29.483

alignment_block:
us-09-598-042a-2 x AA15269 ..
Align seg 1/1 to: AA15269 from: 1 to: 652

265 GGGGTGTACGTGTGAAGGCGCACCAAGCGCTTGGCGGCGCTGAGCGTC.. 312
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
5 GlYArPheIleCysLeuValValThraIleMetAlaThrLeuSerLeuAl 21
313 .....ACTACACCGCTCTGCTGTGATGACATTAGCCAGGGAAG 355
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
21 ArGPrOserPheSerLeuVal..... 28
356 AGACCGTGGGGCCGACAGCTCTGTGGGGGTGAAGAGACCCCGCAGC 405
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
29 .....GlnAspThrThrLeu 33
406 CAGCAGTGGCAGACCGCGCTTCAACAGCGCTTCAAGATGAGGCGCG 455
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
34 GluPrOgluIleAlaProTyrTrpThrAsnThrGlnLysMetGluysAr 50
456 GGTGATGCAAGCGCGCGGTGAGTCTCGTGGGCTCAAGTGGCGCA 505
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
50 glEunIAlaValAlrProAlaIAsnThrValLysPheArGysProAlaG 67
506 GCGGACCGCTCGCGCGACATCAGTGGATGAAAGAGACCGCGCTG 555
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67 LysIAsnProMetLrPheThrMetArgThrLeuLysAsnGlyLysGluPhe 83
556 ACG.....CGCCAGAGCGCGCTGAGCCAGCAAGAAAGTGGAC 596
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
84 LysGlnGlnIAsnArgIleGlyGlyTyrLysValArgAsnGlnIAsnTrpse 100
597 ACTGAGCTGTGAAGACCTCGCGCGGAGACAGCGGCAATACACTGCC 646
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
100 rLeuIleMetGlnSerValValProSerAspLysGlyAsnTyrThrCysv 117
647 GCGTGTGAACCGCGCGCGCATCAAGCGCACCTCAAGTGGATGT 696
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
117 alValGlnAsnGluTyrGlySerThrLeuAsnIAsnThrLysIAsnVal 133
697 ATCCAGCGGACCGCGTTCAGCGCGCTGTCACAGGACGACCGCGTAA 746
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
134 ValGlnArgSerProIAsnArgProIleuGlnAlaGlyLeuProAlaAs 150
747 CAGCAGGTGACTTGGGGGACCAAGCTCTTCAAGTGGAGTGGCA 796
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
150 nAlaserThrValValGlyAspValGluPheValCysLysValTyrS 167
797 GCGACGTGAAGCGGTGATCCAGTGGTGAAGCGCGTGGAG..... 837
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
167 eArPAlaGlnProIAsnIleGlnThrLysIAsnValGlnLysAsnGly 183
838 .....TACGCGCGGAGGCGCGCACACTCCACATGATGTGGCGG 881
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
184 SerLysTyrGlyProAspGly..... 190
882 CCAGAGTTGTGTGCTGCCACAGGTGACGTG.....TGTGCGGCGC 925
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
191 .....LeuProTyrLeuLysValLeuLysHisSerIlyI 202
926 CCGAGCGCTCTACCTCAATAGCTGCTCATACCGCGCGCGCGGAGAC 975
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
202 leAsnSerAsnAlaGlnValAlaLeuPheAsnValThrGlnAla 218
976 GATCGGGCATGTACATCTGCTTGGCCCAACACATGGGCTACAGCTT 1025
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
219 Aspalactylgltutrylleacylsvalserasntryllelglnalaas 235

```





```

816 CCAAGTGGCTGAAGCGCGTGAAGTACGGCGCGGAGCGCCGCAACTCA 865
    |||||||.....:
196 eGlnTrpLeuLysHisIleGluVal.....AsnGlySerL 208
    |||
866 CCATCGATGTGGGCGCCAGAGTTGTG...GTGCTGCCGCGGGGTGAC 912
    |||
208 ySileGlyProAspAsnLeuProGlyValGlnIleLeuLysThrAlaGly 224
    |||
913 GTGAGTGGCGCGCGCGCGGCTCTACTCAATAGCTGCACACACCG 962
    |||
225 Val.....AsnThrThrAspLysGlnMetGluValLeuHisLeuArgAs 239
    |||
963 TGCCCGCGGAGGAGCGGCGGATGATCATCTGCTGGCGCCAAACACA 1012
    |||
239 nValSerPheGlnAspAlaGlyGluTyrThrCysLeuAlaGlyAsnSerI 256
    |||
1013 TGGGCTACAGCTTCCGACGCGCTTCTCACACCGTGTGCCA.....GAC 1056
    |||
256 IeGlyLeuSerHisHisSerIleThrLeuThrValLeuGluAlaLeuGlu 272
    |||
1057 CCAAAACCGCGAGGCGCCACTGTGCGCTCTGCTGCGCCACTAGCT 1106
    |||
273 GluArgPro.....AlaValMetHisSer.ProLeuTyrL 284
    |||
1107 GCCGTGGCGCGGTGTATCGGCATCCAGCGGCGGTGTTCATCTGTG 1156
    |||
284 euGluGly.SerGlySer..... 289
    |||
1157 GCACCTCTCTCTGTGTGTGCGAGCGCCAGAAAGACCTGTGACCC 1206
    |||
290 .....ProGlyLeuGlnGlu.....ProL 296
    |||
1207 GCGCGTGGCGCTCCCTGCTGCGACCGCGCGGAGCGCGCGCA 1256
    |||
296 ySerCysAspLysThrHisThrCysProProCys.....ProAla 309
    |||
1257 CCGCAGCGGAGACAGACCTTCCCTC 1283
    |||
310 ProGluLeuLeuGlyGlyProSerVal 318
seq_name: /STD1/gcgdata/geneseq/geneseq/AA1998.DAT:AAW5721
seq_documentation_block:
ID AAW5721 standard; Protein; 472 AA.
AC AAW5721;
XX
XX 07-DEC-1998 (first entry)
XX
XX Fibrinblast growth factor receptor-tpa fusion protein.
XX
XX Fibrinblast growth factor receptor 1; tissue plasminogen activator;
XX FGFRI; tPA; human; fibrinblast growth factor; FGF; thrombosis;
XX thrombocytopenia; ophthalmic disorder; human; therapy.
XX
XX Chimeric - Homo sapiens.
XX
XX MO9839436-A2.
XX
XX 11-SEP-1998.
XX
XX 03-MAR-1998; 98WO-JP00878.
XX
XX 03-MAR-1997; 97US-0040785.
XX
XX (EISA ) EISAI CO LTD.
XX
XX Kalyanaraman R, Kawai T, Zhu H;
XX
XX WPI; 1998-495843/42.
XX
XX N-PSDB; AAV47648.
XX
XX Fibrinblast growth factor mutein and DNA - having reduced receptor

```

```

PT binding and able to bind heparin, useful for treating and regulating
PT heparin-related disorders e.g. thrombosis
XX
XX Example 2; Page 61-63; 71pp: English.
XX
XX This is the amino acid of a fusion protein comprising human
XX fibrinblast growth factor receptor 1 (FGFR1) and tissue plasminogen
XX activator (tPA). FGFR1 and tPA cDNAs were obtained by PCR
XX amplification and were ligated into pUC18. The resulting FGFR1-tPA
XX fusion (see AAV47648) was incorporated into mammalian expression
XX vector pK4K, and the chimeric protein was expressed in transfected
XX BHK cells. It was used in assays for measuring the binding activity
XX of FGF muteins (see AAW5711-20) to FGFR1. Claimed FGF muteins show
XX reduced FGF receptor binding activity but retain the ability to
XX bind heparin. They are used to treat heparin-related disorders,
XX such as excessive bleeding induced by heparin, ophthalmic disorders
XX and heparin-associated thrombocytopenia and thrombosis.
XX
XX SQ Sequence 472 AA:
XX
XX alignment_scores:
XX Quality: 373.00 Length: 401
XX Ratio: 1.727 Gaps: 14
XX Percent Similarity: 53.865 Percent Identity: 28.180
XX
XX alignment_block:
XX US-09-598-042A-2 x AAW5721 ..
XX
XX Align seq 1/1 to: AAW5721 from: 1 to: 472
XX
XX 337 GACATTAGCCCGAGGAGAGAGACCTGGGCGCCGACAGCTCTGTGGGG 386
XX ||| |||.....:
XX 31 AspAlaLeuProSerSerGluAspAspAspAspAspSerSerSer 47
XX 387 TCAGAG.....GACCCCGCAGCCAGTGGCGACGACCGGCT 427
XX ||| |||.....:
XX 47 rGluGlnLysGluThrAspAsnThrLysProAsnProValAlaProTyrT 64
XX |||
XX 428 TCACAGACCCCTCCAGATGAGGCGCGGCTGATCGACGCGCGGTGGGT 477
XX |||
XX 64 rPThrSerProGluLysMetGluLysLysLeuHisAlaValProAlaAla 80
XX |||
XX 478 AGCTCGTGGCGGCTCAAGTGCCTGGCGGCGACCGCTGGCCGACAT 527
XX |||
XX 81 LysThrValLysPheLysCysProSerSerGlyThrProAsnProThrLe 97
XX |||
XX 528 CAGCTGATGAGGAGGAGCGAGGCTTGACGCGCCGAG..... 567
XX |||
XX 97 uArgTrpLeuLysAsnGlyLysGluPhe...LysProAspHisArgLleG 113
XX |||
XX 568 .GCCGTGAGCCCGAGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGA 615
XX |||
XX 113 LysGlyTyrLysValArgTyrAlaThrTrpSerIleIleMetAspSerVal 129
XX |||
XX 616 CGCGCGGAGGAGCGGCAATACACTGCGCGGTGTCGACACCGCGGG 665
XX |||
XX 130 ValProSerAspLysGlyAsnTyrThrCysIleValGlnLysnGluTyrG 146
XX |||
XX 666 CGCATCAAGCGCCACTACAGAGTGAAGTGAATGATCAAGGAGCGGTTCCA 715
XX |||
XX 146 ySerIleAsnHisThrTyrGlnLeuAspValValGluArgSerProHisA 163
XX |||
XX 163 rGProIleLeuGlnAlaGlyLeuProAlaAsnLysThrValAlaLeuGly 179
XX |||
XX 766 GAGACACAGTCTCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 815
XX |||
XX 180 SerAsnValGlnPheMetCysLysValTyrSerAspProGlnProHisI 196
XX |||
XX 816 CCAAGTGGCTGAAGCGCGTGAAGTACGGCGCGGAGCGCCGCAACTCA 865
XX |||

```



```

130 ValProSerAspLysGlyAsnTyrThrCysIleValGluAsnGluTyrG1 146
666 CGCCATACAGCCACCTACAAAGTGGATGTGATCCAGCGACCCGTTCCA 715
146 YserIleAsnHisThrTyrGlnLeuAspValValGluArgSerProHisA 163
716 AGCCGCTGCTACAGCAGCGACCCGTAACAGCAGCGTGTGAGG 765
163 rgrProlIleGlnIleGlyLeuProIleAsnLysThrValAlaLeuGly 179
766 GGAACACAGCTCTTCAGTGCAGGTGCGCAGCAGCTGAAGCCGTGAT 815
180 SerAsnValGluPheMetCysLysValTyrSerAspProGlnProHisI1 196
816 CCAATGGCTGAAGCGCGTGGAGTACGGCGCCGAGCGCCCAACACATCCA 865
196 eGlnTrpLeuLysHisIleGluVal.....AsnGlySerL 208
866 CCATCGATGTGGCGCCAGAAATTGNG...GTGTCGCCAGCGGTGAC 912
208 yslIeGlyProAspAsnLeuProTyrValGlnIleLeuLysThrAlaGly 224
913 GTGTGTGTCGCGCCGCGAGCGCTCTACCTCAATAAGCTGTATCACCCG 962
225 Val.....AsnThrThrAspLysGluMetAspValLeuHisLeuArgAs 239
963 TGCCCGCCAGACATGCGGCGATGTACATCTGCTTGGCCCAACACCA 1012
239 nValSerPheGlnAspAlaGlyGluTyrThrCysLeuAlaIleAsnSerI 256
1013 TGGCTACAGCTCCGACGAGCGCTTCTCCACCGTGTGTCGACCCCAAAA 1062
256 IeGlyLeuSerHisHisSerIleAlaTrpLeuThrValLeu...GluAlaLeu 271
1063 CCGCGAGGCGCACCTGTGCTCTCTGCTGCGGCACCTAGCTGCGGTG 1112
272 GluGluArgProAlaValMetThrSer.ProLeuTyrValAspAlaArg 288
1113 GCCGTGTGTCATCGGCATCC..... 1132
288 heProAlaGlyAlaArgSerTyrGlnValIleCysArgAspGluLysThr 304
1133 .....CAGCCGCGCTGTCTTC 1149
305 GlnMetIleTyrGlnGlnHisGlnSerTrpLeuArgProValLeuArgSe 321
1150 ATCC.....TGGGCACTGCTGCTGTGCTGTCGCAAG. 1163
321 rAsnArgValGluTyrCysTrp.....CysAsnSerGlyArgAlaGlnC 336
1184 .....CCGAGAGAACCGTCGACCCGCGCGCTCCCTCCCTCC 1222
336 yshHisSerValProValLysSerCysSerGluProArg..... 348
1223 TGCCTG..... 1228
349 CysPheAsnGlyGlyThrCysGlnGlnAlaLeuTyrPheSerAspPheVa 365
1228 ..... 1228
365 lCysGlnCysProGluGlyPheAlaGlyLysCysCysGluIleAspTrA 382
1229 .....GGCACCGCGCCGCGGAGCGCGCCGCG 1255
382 rGAlaThrCysTyrGlnAspGlnGlyIleSerTyrArgGlyThrTrpSer 398
1266 ACCGACGCGGAGACAAAGACCTTCCCTGTTGGCCGCCCTACGCTGGC 1305
399 ThrAlaGlnSerGlyAlaGlnCysThrAsnTrpAsnSerSerAlaLeuAl 415
1306 C 1306
415 a 415

```

```

seq_name: /SUDSI/gcdata/geneseq/geneseq/AA199.DAT:AAV06393
seq_documentation_block:
ID AAV06393 standard; Protein; 472 AA.
XX
AC AAV06393;
XX
DT 20-SEP-1999 (first entry)
XX
DE Human FGFR1-tpa fusion protein.
XX
KW Fibroblast growth factor receptor 1; tissue plasminogen activator;
KW FGFR1; tpa; fusion protein; antagonist; diabetes; cancer;
KW melanoma; tumour; restenosis; rheumatoid arthritis;
KW dermatological disorder; ophthalmological disorder;
KW diabetic retinopathy; therapy.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Protein 1..284 /note="human FGFR1"
FT Peptide 285..294 /note="linker"
FT FT 295..472 /note="tpa"
FT FT
XX
PN W09929640-A2.
XX
PD 17-JUN-1999.
XX
PE 04-DEC-1998; 98WO-US25789.
XX
PR 15-MAY-1998; 98US-0079343.
PR 05-DEC-1997; 97US-0986248.
XX
PA (EISA ) EISAI CO LTD.
XX
PI Balaji VN, Castillo RS, Chan MF, Kols A, Ramnarayan K;
PI Schove L;
XX
DR WPI: 1999-404912/34.
DR N-PSDB; AAX59330.
XX
PT Composition used to modulate interaction of fibroblast growth factor
PT at receptor
XX
XX
XX Example 28; Page 154-155; 155pp; English.
XX
XX The present sequence represents a fusion protein comprising
XX amino acids 1-284 of human fibroblast growth factor receptor 1
XX (FGFR1), a peptide linker and amino acids 1-178 of human tissue
XX plasminogen activator (tpa). The fusion protein was expressed
XX in mammalian cells using expression vector pK4K following PCR
XX amplification of FGFR1 and tpa cDNAs (see AAX59330). It was used
XX to screen for FGF antagonists that inhibit FGF binding to its
XX receptor. Such antagonists, especially aromatic acids including
XX invention, are useful for treating FGF mediated diseases including
XX diabetes, cancer including melanoma and tumour growth, restenosis,
XX in stent restenosis, rheumatoid arthritis, proliferative
XX dermatological disorders, ophthalmological disorders including
XX corneal clouding following excimer laser surgery, closure of
XX trabeculocornies, hyperproliferation of lens epithelial cells
XX following cataract surgery, the recurrence of pterygia and diabetic
XX retinopathy, and other proliferative diseases.
XX
SQ Sequence 472 AA;

```

alignment\_scores: Quality: 373.00 Length: 401





















1070 GGCCACCTGGCTCTGCTGCGCCACCTAGCTGCGTGGCCGCTG 1119  
 248 .....AlaValMethIser.ProLeuTyrIleuGluGly.SerG 260  
 1120 GTCATCGGCATCCAGCGCGGTGTCTTCAATCGGACCCCTGCTCT 1169  
 260 LysSer..... 261  
 1170 GTGCTTTGCCAGGCCCAAGAAGACCGCTGACCCCGCGCTCCCTC 1219  
 262 .....ProGlyIleuGlnGlu.....ProLysSerCysAspLys 272  
 1220 CCCTGCTGGGACCGCCGCGGGGAGAGCGCCGACCGACCGAGAGAC 1269  
 272 sThIstHrcysProProcys.....ProAlaProGluIleuLeuG 286  
 1270 AAGGACCTTCCTC 1283  
 286 LysGlyProSerVal 290

seq\_name: /STDS1/gcgdata/geneseq/geneseq/AA2000.DAT:AA36232

seq\_documentation\_block:

ID AAB36232 standard; Protein; 304 AA.

AC AAB36232;

DT 19-FEB-2001 (first entry)

DE Rat FGFR-1 flagged protein.

KW Rat; Fas receptor; gene therapy; vascular cell; stenosis; FGFR-1;

KW accelerated graft arteriosclerosis; vascular occlusive disease.

OS Rattus sp.

OS Synthetic.

PN WO200063369-A2.

PD 26-OCT-2000.

PF 18-APR-2000; 2000WO-EP03532.

PR 19-APR-1999; 99US-0294176.

PR 10-MAY-1999; 99US-0307690.

PA (NOVS ) NOVARTIS AG.

PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.

PA (UYJO ) UNIV JOHNS HOPKINS.

PI Ballermann BJ, Goldschmidt P, Movva RN, Sanfillippo A;

DR MPI: 2000-672735/65.

DR N-PSDB; AAC66561.

PT Genetic modification of target cells for inhibiting excessive

PT proliferation, comprises transferring, to cells or progenitors, the DNA

PT sequence encoding Fas antigen or receptor, or the fibroblast growth

PT factor receptor -

PS Example 4; Page 40-41; 41pp; English.

CC The present invention describes a method of gene therapy which involves

CC the use of Rasr or FGFR-1 coding sequences to treat vascular diseases.

CC The invention also provides vectors for use in this method. The treatable

CC diseases include vascular occlusive diseases associated with cell

CC proliferation, such as accelerated graft arteriosclerosis and other forms

CC of stenosis.

CC Sequence 304 AA;

alignment\_scores:  
 Quality: 367.00 Length: 310  
 Ratio: 1.963 Gaps: 9  
 Percent Similarity: 60.323 Percent Identity: 30.323

alignment\_block:

US-09-598-042a-2 x AAB36232 ..

Align seg 1/1 to: AAB36232 from: 1 to: 304

147 GGACCCGCGCGCTGACACATGTGGACCAAGATGGCCGACATCCACA 196  
 11 GlyAlaValAlaAspTyrLysAspAsp.....AspAs 22  
 197 GCGGCTGGAGCCGCTCCGCGTGTGCGGAGGGGCTGAAGTGAAGCAG 246  
 22 pLysLeuAlaGProAlaPro..ThrLeuProGlu..... 32  
 247 GTGAGGCGGAGATGCCGCGGTGTACGTGTGCAAGGCCACCAAGGCTT 296  
 32 ..... 32  
 297 CGGACCGCTGAGCTGACACACCTCGTGTGCTGATGACATTAGCC 346  
 33 .....GlnAspAlaLeuP 37  
 347 CAGGAGAGAGAGCTGGGGCCGACAGCTCTGTGGGGTCAAG... 393  
 37 roSerSerGluAspAspAspAspAspAspSerSerSerGluGluLys 53  
 394 .....GACCCGCGCAGCCAGCAGTGGGGCAGACCGCGCTTACACAGCC 437  
 54 GluThrAspAsnThrLysProAsnProValAlaProTyrTrpThrSerPr 70  
 438 CTCGAGATGAGGCGCGGGGTGATGACAGCGCGCGGTGAGTCCGTGC 487  
 70 OGluLysMetGluLysLysLeuHisAlaValProAlaAlaLysThrValL 87  
 488 GGCTCAAGTGTGCGCCAGCGGCGACCTCGGCCGACATCAGCTGAGAG 537  
 87 YspHelysCysProSerSerGlyThrProSerProThrLeuAlaGlyTrpLeu 103  
 538 AAGGAGCAGCAGCCCTTGACCGCCACAG.....GCCGCTCA 575  
 104 LysAsnGlyLysGluPhe..LysProAspHisArgIleGlyLysTrpLys 119  
 576 GCCCAGGAGAGAGAGTGCAGACTGAGCTGAGAGACCTGCGCGGAGG 625  
 119 SValArgTyrAlaThrTrpSerIleIleMetAspSerValValProSerA 136  
 626 ACAAGCGCAATACACCTGCGCGGTGTGCAACCGCGCGGCGCATCAAC 675  
 136 sPlysGlyAsnTyrThrCysIleValGluAsnGluTyrGlySerIleAsn 152  
 676 GCCACCTACAGAGTGTGATGATCCAGCGGACCCGTTCCAGCCGCTGT 725  
 153 HisThrTyrGluLeuAspValValGluArgSerProHisArgProIleLe 169  
 726 CACAGGACGACCCCGCTGAACAGAGGTGACTTCGAGGGGACACAGT 775  
 169 uGlnAlaGlyLeuProAlaAsnLysThrValAlaLeuGlySerAsnValG 186  
 776 CTTTCAGTCAAGTGTGCGAGCGAGCTGAAGCGGATGATCAAGTGGCTG 825  
 186 IupMetCysLysValTyrSerAspProGlnProHisIleGlnTrpLeu 202  
 826 AAGCGCGTGTGAGTACGGCGCGGCGGCGGCGGACACACATCCACATG 875  
 203 LysHisIleGluVal.....AsnGlySerLysIleGlyTrp 214  
 876 GGGCGGCGCAAGAGTTGTG...GTGCTGCCACGCGGTGAGCTGTGCGC 922  
 214 OAspAsnLeuProTyrValGlnIleLeuLysThrAlaGlyVal.....A 229

```
923 GAGCGGAGCGCTCTACATCATAGCTGTCTATCATCCGCGCCGCGCAG 972
    ::::::::::: ||| :::::::::::
229 srthThrAspLysGluMetGluValLeuHisLeuArgAsnValSerPhe 245
973 GAGCATGCGCGCATGTACATCTGTGGCGGCGCACACCATGGCTACAG 1022
    ::::::::::: ||| ||::::::::: ||::::::::: |||
246 GluAspAlaGlyGluTyrThrCysLeuAlaGlyAsnSerIleGlyLeuSe 262
1023 CTTCGCGACGCGCTCTCTACCGCTGCTG 1050
    | ::::::::::: ||::::::::: ||
262 rHsHsSerAlaTrpLeuThrValLeu 271

seq_name: /SISL/gcdata/geneseq/geneseq/AA2000.DAT:AA97175
seq_documentation_block:
ID   AA97175 standard; Protein: 488 AA.
XX
AC   AA97175;
XX
DT   04-DEC-2000 (first entry)
XX
DE   Human FGF-R1 Extracellular domain-Ig Fc fusion protein 6.
XX
KW   FGF-R; fibroblast growth factor receptor; extracellular domain; IgG1;
KW   immunoglobulin; Gl; oligomerization domain; Fc region; fusion protein;
KW   inhibitor; dimer; antagonist; cytostatic; anti-diabetic; vulnery;
KW   ophthalmological; anti-proliferative.
XX
OS   Homo sapiens.
XX
FH   Key
FT   Peptide
    1..21
    /label= FGF-R1-signal-peptide
FT   Domain
    22..257
    /label= FGF-R1-extracellular domain
    /note= "The Ig I segment and acid box are deleted"
FT   Domain
    59..111
    /label= Ig_II-segment
FT   Domain
    157..222
    /label= Ig_III-segment
FT   Region
    258..488
    /label= Human_IgG1_Fc-region
    /note= "Contains hinge region and domains CH2 and CH3"
FT   Misc-difference 276
    /label= L276E
    /note= "This mutation decreases the affinity of the Fc
FT   portion for Fc receptors"
FT   Misc-difference 376
    /label= p376S
    /note= "This mutation decreases the affinity of the Fc
FT   portion for complement"
XX
PN   WO200046380-A2.
XX
PD   10-AUG-2000.
XX
PF   07-FEB-2000; 2000WO-US03166.
XX
PR   08-FEB-1999; 99US-0119002.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Kavanaugh WM, Ballinger M;
XX
DR   WPI: 2000-514961/46.
XX
DR   N-PSDB: AAA52132.
XX
PT   New polypeptide comprising a fibroblast growth factor receptor
PT   extracellular domain fused to a heterologous oligomerization domain for
PT   treating FGF-, angiogenesis-, or FGF receptor-mediated disorders
XX
PS   Claim 14; Page 68-69; 70pp; English.
```

```
XX
CC   Novel fusion protein constructs comprise a fibroblast growth factor (FGF)
CC   receptor (FGF-R) extracellular domain (ECD) lacking the immunoglobulin
CC   (Ig) I segment fused to a heterologous oligomerization domain that
CC   comprises an immunoglobulin Fc region, hinge region, CH1, CH2, CH3 or CH4
CC   region, or light chain of an immunoglobulin molecule, or a peptide with a
CC   leucine zipper motif. The Ig I segment is not necessary for binding of
CC   acidic FGF and basic FGF (bFGF). The Ig I deletion further increases the
CC   affinity for aFGF and heparin, protects the core of the molecule from
CC   proteolysis, and abrogates the heparin requirement for aFGF binding. The
CC   new fusion polypeptides are better FGF inhibitors than FGF-R monomer
CC   proteins. The FGF-R-Ig Fc fusion dimers are active as FGF antagonists at
CC   subnanomolar concentrations and were 20-fold more potent than the FGF-R
CC   monomer protein as competitors of bFGF binding to immobilized FGF-Rs. The
CC   fusion constructs are useful to treat FGF-, angiogenesis-, or
CC   FGF-R-mediated disorders, such as tumourigenesis (e.g. bladder, breast,
CC   lung, rectal, testis and cervical tumours), neovascularization (e.g.
CC   diabetic retinopathy, neovascular glaucoma, wound healing and corneal
CC   scarring) and hyper-proliferation of vascular smooth muscle cells (e.g.
CC   postangioplasty and postatherectomy restenosis).
CC
SQ   Sequence 488 AA;
XX

alignment_scores:
    Quality: 366.50      Length: 289
    Ratio: 2.118        Gaps: 10
    Percent Similarity: 59.862      Percent Identity: 31.834

alignment_block:
US-09-598-042a-2 x AA97175 ..

Align seg 1/1 to: AA97175 from: 1 to: 488

391 GAGGACCCGCGCAGCGAGTGCGGCGCAGCGCGCTGCACAGCCCTC 440
    ||::||::||::||::|||
29 GluGlnProValAlaProTyrTrp.....ThSerProI 40
441 CAAGATGAGCGCGCGGTGATCGCACGCGCGCTGGTACCTCGTCGGCG 490
    ||::||::||::||::|||
40 uLysMetGluLysLysLeuHisAlaValProAlaAlaLysThrValLysP 57
491 TCAGTGCCTGGCGCACCGGCGACCCCTCGCGCGCGACATCAGTGAAG 540
    ::||::||| ::||::||| ::||::|||
57 heLysCysProSerSerGlyThrProAsnProThrLeuArgTrpLeuLys 73
541 GAGGACCGCGCTGTGACGCGCGCAGAG.....GCGCGTCGACC 578
    :::: :::: ::||::||::
74 AsnGlyLysGluPhe...LysProAspHisArgIleGlyGlyTyrLysVa 89
579 CAGGAGAGAGAGTGGACACTGAGCCTGAGAACCTGCGCGCGAGAGACA 628
    ||| ::::||| ::::||| ::::|||
89 LArgTyrAlaThrTrpSerIleIleMetAspSerValAlaProSerAspL 106
629 GCGGCAAAATACACTGCGCGCGGTGTGACACCGCGCGCGCATCAAGCC 678
    ::||::||::||::||| ::||::||| ::||::|||
106 ysgLysAsnTyrThrCysIleValGluAsnGlnTyrGlySerIleAsnHis 122
679 ACCTACAGAGTGATGTGATCCAGCGACCGCTTCACAGCGCGGTGCTAC 728
    ||::||::||::||::||| ::||::||| ::||::|||
123 ThrTyrGlnLeuAspValValGluArgSerProHisArgProIleLeuG 139
729 AGGACAGCACCCGCTGACACAGCAGGTGACTCGGGGAGACACGCTCT 778
    :::: ||::||| ||::||| ::||::|||
139 nAlaGlyLeuProAlaAsnLysThrValAlaLeuGlySerAsnValGluP 156
779 TCCAGTGCAGAGTGGCGCGCAGCGATGAACCGGTATCCAGTGCAGAG 828
    ||::||::||::||| ||::||| ::||| ::|||
156 hewetCysLysValTyrSerAspProGlnProHisIleGlnTrpLeuLys 172
829 CGGCTGAGTACGCGCGCAGGCGCGCACACACTCCACATCGATGTGGG 878
    ::::||| ::::||| |||
173 HisIleGluVal.....AsnGlySerLysIleGlyProAs 184
```







**THIS PAGE IS BLANK**







```

108 GCTGGCCGCGCATGAGGCTGAGTCCGACAGTGGAGGGAGCCGCGC 157
109 |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
41 AlaGlyProHisCysAlaAlaAlaVal.ProValGlyGlyAspProP 57
158 CGCTGACCATGTGACCAAGAGATGGCCGACACATCCACAGCGCTGAGC 207
159 |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
57 rGleuThrMetTrpThrLysAspLysArgThrIleHisSerGlyTrpSer 73
208 CGCTCCGCGGTGCTGGCCGACAGGGCTAAGTGAAGAGGTGGAGCGGGA 257
209 |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
74 ArgPheArgValLeuProGlnGlyLeuLysValGlnValGlnArgG 90
258 GGAATGGCGGCTGATGCTGACAGGCGCCACCAAGCGCTGGCGCGCTGA 307
259 |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
90 uAspAlaGlyValIleValIleCysLysAlaThrAsnGlyPheGlySerLeu 107
308 GCGTCAACTACACCTCGCTGCTGTGATGACATTAAGCCAGGAGAGAG 357
309 |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
107 eValAsnIleValIleValIleValLeuAspIleSerProGlyLysGlu 123
358 AGCCTGGGGCGCGACGCTCCTGCGGGGTCAGAGAGAGCCCGCGCAGCA 407
359 |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
124 SerLeuGlyProAspSerSerSerGlyGlyGlnGlnAspProAlaSerG 140
408 GCACTGGGCGACAGCGCGCTTCACACAGCCCTCCAGATGAGCGCGCGG 457
409 |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
140 nGlnTrpAlaArgProAlaArgPheThrGlnProSerLysMetArgArgArg 157
458 TGATCGCAGCGCGCGTGGTACGCTCCGCTGCGGCTCAAGTGGCTGGCC 507
459 |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
157 AlIleAlaArgProValIleGlySerSerValArgLeuLysCysValAlaSer 173
508 GGGCACCCTCGCGCGCGCATCATCAGTGTGATGAGAGAGAGAGCGCTGAC 557
509 |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
174 GlyIleProArgProAlaArgPheThrGlnProSerLysMetArgArgArg 190
558 GCGCCGAGAGCGCGCTGAGCCCGACAGAGAGAGAGAGTGAACACTGAG 607
559 |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
190 rArgProGlnAlaIleAlaGlnProArgLysLysLysTrpThrLeuSerLeu 207
608 AGAACCTGGCGCGGAGAGAGAGCGGAAATACACTGCGCGCTGCGAAC 657
609 |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
207 yAsnLeuAlaArgProGlnAspSerLysLysLysTrpCysArgValSerAsn 223
658 GCGCGGCGCGCATCAAGCGCCACTACAGAGTGTGATCCAGCGGAC 707
659 |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
224 ArgAlaGlyAlaIleAsnAlaThrTrpLysValAlaIleGlnArgTh 240
708 CCGTTCACAGCGCGTGTGACAGGACAGCAGCCCGTGAACACGAGGTGG 757
709 |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
240 rArgSerLysProValLeuThrGlyThrHisProValAsnThrThrValA 257
758 ACTTCGGGGGAGACAGCTCTCCAGTGAAGTGGCAGAGAGAGTGAAG 807
759 |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
257 sprPheGlyGlyThrThrSerPheGlnCysLysValArgSerAspValLys 273
808 CCGGTGATCAGTGGCTGAAGCGCGGTGAGTACGCGCGCGAGCGCGCCA 857
809 |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
274 ProValIleGlnTrpLeuLysArgValGlnIleGlyLysGlnArgGln 290
858 CAACATCAACCATCATGTGGCGCGCGCGAGAGTGTGCTGCGCGCAGG 907
859 |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
290 sAsnSerThrIleAspValGlyGlyGlnLysPheValValLeuProThrG 307
908 GTGACGTGTGCTGCGCGCGCGCGCGCTCTACCTCAATAAGCTGCTC 954
909 |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
307 LysAspValIleTrpSerArgProAspGlySerTrpLeuAsnLysProLeu 322

```

seq\_name: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:US-09-383-586-30

seq\_documentation\_block:  
; Sequence 30, Application US/09383586

```

; Patent No. 6242419
; GENERAL INFORMATION:
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Abernethy, Neva
; APPLICANT: Onrust, Rene
; APPLICANT: Kumble, Anand
; APPLICANT: Murlison, Greg
; TITLE OF INVENTION: Compounds isolated from stromal cells
; FILE REFERENCE: 11000.1037c1
; CURRENT APPLICATION NUMBER: US/09/383.586
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 30
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Mouse
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (123)...(123)
US-09-383-586-30

alignment_scores:
  Quality: 645.50      Length: 127
  Ratio: 5.164        Gaps: 1
  Percent Similarity: 98.425  Percent Identity: 95.276

alignment_block:
US-09-598-042a-2 x US-09-383-586-30 ..

Align seg 1/1 to: US-09-383-586-30 from: 1 to: 126

754 GTGACCTGGGGGAGCCAGCGCTTCACAGTGTGATGAGAGAGAGTGGAGCT 803
1 ValAspPheGlyGlyThrThrSerPheGlnCysLysValArgSerAspVal 17
804 GAAGCGGTGATCACTGCTGAGCGCGGTGAAGCGCGTGAAGTACGGCGCGAGGCC 853
17 LysProValIleGlnTrpLeuLysArgValGlnIleGlySerGlnGlyA 34
854 GCCACAACTCCACCATCGATGTGGCGCGCGCCAGAGATTGTGTGCTGCGCC 903
34 rGlnAsnSerThrIleAspValGlyGlyGlnLysPheValValLeuPro 50
904 ACGGTGACGTGTGTCGCGCGCGCGAGCGGCTCTACCTCATTAAGCTGCT 953
51 ThrGlyAspValIleTrpSerArgProAspGlySerTrpLeuAsnLysLeu 67
954 CATCACCGCTGCCCGCGAGAGAGATGCGGGCATGATCATCTGCTTGGCG 1003
67 uIleSerArgAlaArgGlnAspAspAlaGlyMetTrpIleCysLeuGlyA 84
1004 CCAACACCATGAGGCTACAGCTTCCGAGCGCGCTTCCATCAACCGTGCGCA 1053
84 LAsnThrMetGlyTrpSerPheArgSerAlaPheLeuThrValLeuPro 100
1054 GACCCAAACCGCGAGAGCGCACCTGTGAGCGCTCTCTGCTGCGCGCACTAG 1103
101 AspProLysProProGlyProPheMetAlaSerSerSerSerSerThrse 117
1104 CTTGCGGTGGCGCGGTGCTCATCGCATCCCA 1134
117 rLeuProTrpPro...Val**GlyIlePro 126

```

seq\_name: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:US-08-451-822A-13

seq\_documentation\_block:  
; Sequence 13, Application US/08451822A  
; Patent No. 5863888  
; GENERAL INFORMATION:

```
APPLICANT: Dionne, Craig A
APPLICANT: Crumley, Greg
APPLICANT: Jeye, Michael C
APPLICANT: Schlessinger, Joseph
TITLE OF INVENTION: Fibroblast Growth Factor Receptors
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESS: Rhone-Poulenc Rorer Legal Department
STREET: 500 Arcola Road
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451,822A
FILING DATE: 26-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/323,430
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/934,372
FILING DATE: 21-AUG-1992
APPLICATION NUMBER: US 07/549,587
FILING DATE: 06-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky, Martin
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: A0496E
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3816
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 821 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-451-822A-13

alignment_scores:
Quality: 462.00 Length: 387
Ratio: 1.958 Gaps: 10
Percent Similarity: 60.982 Percent Identity: 31.008

alignment_block:
US-09-598-042a-2 x US-08-451-822a-13 ..
Align seg 1/1 to: US-08-451-822a-13 from: 1 to: 821

100 GTGCCCGCGGTGGCGCGCATATGCGGCTGCAGTCCCGCATGGAGGGGGA 149
||||| |
51 ValAlaIaIaProGlyGluSerLeuGluValAlaIaIaIaIaIaIaIaIaIa 65
150 CCCGCCCGCGGTGCACATGTGGACCAAGATGCGCGCACCATCCACAGCG 199
||| |
66 AspAlaIaIaValIleSerTrpThrIysAspGly.....ValHisLeuG 80
200 GCTGAGCGCGCTTCCGCTGCTG...CCGAGGGGCTGAAGTGAAGCAG 246
||| |
80 LysProAsnAsnArgThrValLeuIleGlyGluTyrLeuGlnIleLysGly 96
247 GTGAGGAGGAGATGCGCGGTGTACGTGTGCAAGGCGCACCAAGCGCTT 296
||| |
97 AlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 113
```

```
297 CGGAGCCTGAGCCGTCACATACACCCCTGCTGCTGATGACATTAGC 346
||| |
113 LAspSerGluThrTrpTyrPheMetValAsnValHisAspAlaIleSer 130
347 CAGGAAAGAGAGAGCCTGGGGCCGACAGCTCTGCGGGCAAGAGAGAC 396
||| |
130 ergLysAspAsp.....GluAspAspThrAspGlyAlaGluAsp 142
397 CCCGCCAGCCAG.....CAGTGGGACAGCCGCGCTTCACACAGCC 437
||| |
143 PheValSerGluAsnSerAsnIysArgAlaProTyrTrpThrAsnTh 159
438 CTCCAAGATGAGCCCGCGGTGATCGCACGCGCCCTGGGTAGCTCCGTC 487
||| |
159 rGluLysMetGluLysArgLeuHisAlaValProAlaIaIaIaIaIaIa 176
488 GAGTCAAAGTGGCGCGGAGCGGACCTCGGCGGACATACGTAAGTGA 537
||| |
176 yspPheArgCysProAlaGlyGlyAsnProMetProThrMetArgTrpLeu 192
538 AAGGACAGCAGGCGCTTGACG.....CGCCAGAGCGCGCTGAGCC 578
||| |
193 LysAsnGlyLysGluPheLysGlnGluHisArgIleGlyGlyTyrLysVa 209
579 CAGGAAAGAGAGTGGACACTGAGCTTGAAGACCTGCGCGCGGAGAGCA 628
||| |
209 LArgAsnGlnHisTrpSerLeuIleMetGluSerValAlaProSerAspL 226
629 GCGGCAATACACCTGCGCGGTGACGCGGCGGAGCGGCGGCGCATACGCC 678
||| |
226 ysgLysAsnTyrThrCysValAlaGluAsnGluTyrGlySerIleAsnHis 242
679 ACCTACAGGTGATGTGATCCAGCGGACCCGTTCCAGCCCGTCTAC 728
||| |
243 ThrTrpHisLeuAspValValGluArgSerProHisAspProIleLeuG 259
729 AGCAGCAGCAGCCCGTGAACAGCAGCGGTGACTCGGGGGAGCAGCTCT 778
||| |
259 nAlaGlyLeuProAlaAsnAlaSerThrValValGlyLysPheValGluP 276
779 TTCAGTGCAGAGTGGCGGAGCGGAGCGGATGACGAGTGGTGAAG 828
||| |
276 heValCysLysValTyrSerAspAlaGlnProHisIleGlnTrpLys 292
829 CGCGTGAAG.....TACGCGCGGAGGCGCGCCGACCAACATC 863
||| |
293 HisValGluLysAsnGlySerLysTyrGlyProAspGly..... 305
864 CACCATCGATGTGGCGGCGCAGAAAGTTGTGATGCTGCCACGCGTACG 913
||| |
306 .....LeuProTyrLeuLysV 311
914 TGTGTCGCGCGCGCGAGCGCTCTACCTCATTAAG.....CTG 951
||| |
311 alLeuLysAlaIaIaGlyValAsnThrThrAspLysGluIleGluValLeu 327
952 CTCATACACCGCTGCGCGGAGCAGATGCGGAGATGATCATCTGCTGG 1001
||| |
328 TyrIleArgAsnValThrPheGluAspAlaGlyGluTyrTrpCysLeuAl 344
1002 CGCCAAACACATGGGATACAGCTTCCGAGCGCTTCTCTCAACCGTGTG 1051
||| |
344 agLysAsnSerIleGlyIleSerPheHisSerAlaIaIaIaIaIaIaIa 361
1052 CAGAACCCAAACCGCAGGCGCAGCTGTGGGCTCTGCTCGTGGGCACT 1101
||| |
361 roAlaPro.....GlyArgGluLysGluIleThr 370
1102 AGCCTGCGCGTGGCGCGGTGATGAGGATCCAGCGCGGCTGTCTTCAT 1151
||| |
371 AlaSerProAspTyrLeuGluIleAlaIaIaIaIaIaIaIaIaIaIaIa 387
```

1152 CTTGGGACCTGCTGCTGCTTTCGACGACCCAGAGAGCCGTGCA 1201  
387 uillealacymetvalthrvalilleucysargmetlysasnhr 404  
1202 CCCCCGGGCGCT 1212  
404 hrltyslpspro 407

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:us-08-471-570-6

seq\_documentation\_block:

Sequence 6, Application US/08471570  
Patent No. 5750371  
GENERAL INFORMATION:  
APPLICANT: IGARASHI, KOICHI  
APPLICANT: SENOO, Masaharu  
APPLICANT: MATANABE, Tatsuya  
TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &  
ADDRESSEE: CUSHMAN  
STREET: 130 Water Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: US  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,570  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/149,664  
FILING DATE:  
APPLICATION NUMBER: US 07/743369  
FILING DATE: 16-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: LINEK, Ernest V  
REGISTRATION NUMBER: 29622  
REFERENCE/DOCKET NUMBER: 40897  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)523-3400  
TELEFAX: (617)523-6440  
TELEX: 200291 STR UR  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 643 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-471-570-6

alignment\_scores:

Quality: 452.50 Length: 385  
Ratio: 1.926 Gaps: 10  
Percent Similarity: 61.039 Percent Identity: 30.909

alignment\_block:

US-09-598-042A-2 x US-08-471-570-6 ..

Align seg 1/1 to: US-08-471-570-6 from: 1 to: 643

100 GTGGCCCGGCTGGCCGACATATGCGCTGACGCGCAAGTGGAGGGA 149  
||||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
51 ValAlaIaIaProGlyIuSerIeuGluValArgCysIeuLeuLys..... 65  
150 CCGCGCGCGCTGACCATGTGACCAAGATGGCGCGCACCATCCACAGCG 199

66 AspaIaIaValIleSerTrpTrpLysAspGly.....ValHisIeuG 80  
200 GCTGGAGCCGCTTCGCGTGTCTG...CCGAGGGGCTGAAGGTGAAGACAG 246  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
80 LyrAsnAsnArgThrValIleuIleGlyTrpIleuGlnIleLysGly 96  
247 GTGGAGCGGAGGATGCCGCGTGTACGTGTGCAAGCCACCAAGCGCTT 296  
97 AlaThrProArgAspSerGlyLeuTrpAlaCysThrAlaSerArgThrVa 113  
297 CGGACGCTGAGCGTCACTACACCCCTGCTGCTGTGATGATTAAGCC 346  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
113 LAspSerGluThrTrpTrpPheMetValAsnValThrAspAlaIleSer 130  
347 CAGGAGAGAGAGACCCCTGGGCGCGACAGCTCCCTGTGGGGGTCAAGAGAC 396  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
130 eGlyAspAsp.....GluAspAspTrpAspGlyAlaGluAsp 142  
397 CCGGCGACGCGAG.....CAGTGGGACAGCGCGGCTTCACACAGCC 437  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
143 PheValSerIuAsnSerIAsnAsnLysArgAlaProGlyTrpTrpHisnTh 159  
438 CTCGACAGATGAGCGCGCGGTGATCGACGCGCGCGGTGAGTCCGTGC 487  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
159 rGlyLysMetGluLysArgLeuHisAlaValProAlaIaAsnThrVal 176  
488 GGCTCAAGTGGCGGCGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 537  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
176 ySpHeArgCysProIaGlyLysProMetProThrMetArgTrpLeu 192  
538 AAGGAGGACGCGCGCTTGACG.....CGCCAGAGGCGCGCTGAGCC 578  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
193 LysAsnGlyLysGluPheLysGlnIuHisArgIleGlyLysLysVa 209  
579 CAGGAGAGAGAGTGGACACTGACCTGAAGAACTGGCGCGGAGAGACA 628  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
209 LArgAsnGlnHisTrpSerLeuIleMetGluSerValIaProSerAsp 226  
629 GCGCAAAATGACCTGCGCGCGGTGTCGAAACGCGCGCGCGCGCGCGCGCG 678  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
226 yGlyAsnIuTrpTrpCysValIaGluAsnGluTrpGlySerIleAsnHis 242  
679 ACCTAAGGTGATGATGATCAGCGAGCCGCTTCCAGAGCCGCTGTAC 728  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
243 ThrTyHisIeuAspValIaGluArgSerProHisArgProIleuGlu 259  
729 AGGACGAGCCGCGCTGAAACACAGCGGTGACTTGGGGGAGACAGTCT 778  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
259 nAlaGlyLeuProAlaAsnAlaSerThrValIaGlyGlyAspValGluP 276  
779 TTCAGTGCAGGTGCGCCAGCGAGCGGTGATCCGAGGTGGCGTGAAG 828  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
276 heValCysLysValIuTrpSerAspAlaGlnProHisIleGlnTrpIleLys 292  
829 CGCGTGGAG.....TACGGCGCGGAGCGCGCGCGCGCGCAACTC 863  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
293 HisValGluLysAsnGlySerLysTrpGlyProAspGly..... 305  
864 CACCATGATGTGGCGCGCAGAAATTGTGTGCTGCCCGACGCGGTGACG 913  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
306 .....LeuProTyLeuLysVal 311  
914 TG.....TGTCGCGCGCGCGAGGCTCTACTCAATAAGTGCATC 957  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
311 AlLeuLysHisSerGlyIleAsnSerSerAsnAlaGluValIaLeu 327  
958 ACCGTCGCCCGCAGAGAGCATGCGGCGCATGATGCTGCTGGCGGCA 1007  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
328 PheAsnValThrGluLysAspAlaGlyGluTrpIleCysLysValSerAs 344  
1008 CACCATGGGCTACAGCTTCGCGAGCGGCTTCTGACGCTGTCGAGACC 1057  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||





```

914 TG.....TGCTCGGCGCCGACGAGGCTCTACCTCATATAGCTGCTCATC 957
      |||      |||      |||      |||      |||      |||      |||
311 alleuylshiserclylleasnserserAsnAlaIuValleuAlaIeu 327
      |||      |||      |||      |||      |||      |||      |||
958 ACCGTGGCCCGACGAGCATGGCGGCGATGTACATCTGCTGGCGCCAA 1007
      |||      |||      |||      |||      |||      |||      |||
328 pheasnValThrGluAlaAspAlaGlyIuIuIuIuIuIuIuIuIuIu 344
      |||      |||      |||      |||      |||      |||      |||
1008 CACCATGGGCTACAGCTTCGCGAGCGCTTCCTCCATCCGTGCTGCACAC 1057
      |||      |||      |||      |||      |||      |||      |||
344 ntyrileGlyAlaAsnGlnserAlaIuIuIuIuIuIuIuIuIuIuIu 361
      |||      |||      |||      |||      |||      |||      |||
1058 CAAACCCCGACGCGCCACCTGTGGCTCTGCTGCTGCGGCATAGCCTG 1107
      |||      |||      |||      |||      |||      |||      |||
361 lnglnAlaIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIu 373
      |||      |||      |||      |||      |||      |||      |||
1108 CCGTGGCGCCGTCATGCGATCGGCGCGGCGCTGTCTCATCCCTGGG 1157
      |||      |||      |||      |||      |||      |||      |||
374 ProAspIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIu 390
      |||      |||      |||      |||      |||      |||      |||
1158 CACCTGCTCTGCTGCTTTCGAGCGCCGACAGAGACCGCTGCACCCCG 1207
      |||      |||      |||      |||      |||      |||      |||
390 acysmetValIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIu 407
      |||      |||      |||      |||      |||      |||      |||
1208 CGCCT 1212
      |||      |||      |||      |||      |||      |||      |||
407 yspPro 408

```

seq\_name: /cgn2\_6/ptodata/2/laa/5a\_COMB.pep:US-08-471-570-14

```

seq_documentation_block:
; Sequence 14, Application US/08471570
; Patent No. 5750371
; GENERAL INFORMATION:
; APPLICANT: IGARASHI, Koichi
; APPLICANT: SEMO, Masaharu
; APPLICANT: MATAMBE, Tatsuya
; TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,570
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/149,664
; FILING DATE:
; APPLICATION NUMBER: US 07/743369
; FILING DATE: 16-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: LINEK, Ernest V
; REGISTRATION NUMBER: 29822
; REFERENCE/DOCKET NUMBER: 40897
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-471-570-14

```

```

alignment_scores:
  Quality: 437.50      Length: 338
  Ratio: 2.093        Gaps: 9
  Percent Similarity: 61.834      Percent Identity: 32.544

```

alignment\_block:

US-09-598-042a-2 x US-08-471-570-14 ..

Align seg 1/1 to: US-08-471-570-14 from: 1 to: 355

```

100 GTGGCCCGCGCTGGGCGCCGACATATCGGCTGCAGTCCGACAGGAGGGA 149
      |||      |||      |||      |||      |||      |||      |||
31 ValAlaIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIu 45
      |||      |||      |||      |||      |||      |||      |||
150 CCGCGCGCGCTGACCATGTGACCAAGATGCGCGACCATCCACACG 199
      |||      |||      |||      |||      |||      |||      |||
46 AspAlaIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIu 60
      |||      |||      |||      |||      |||      |||      |||
200 GCTGGAGCGCTTCGCGCTGCTG...CCGAGGCGCTGAGAGTGAACAG 246
      |||      |||      |||      |||      |||      |||      |||
60 IuProAsnAsnIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIu 76
      |||      |||      |||      |||      |||      |||      |||
247 GTGAGGCGGAGGATGCGCGCTGCTAGTGTGCAAGCGCCACACAGGCTT 296
      |||      |||      |||      |||      |||      |||      |||
77 AlaThrProArgAspSerGlyIuIuIuIuIuIuIuIuIuIuIuIuIu 93
      |||      |||      |||      |||      |||      |||      |||
297 CGGCGCGCTGAGCGCTCACTACACCTCTGCTGCTGATGATGATGATG 346
      |||      |||      |||      |||      |||      |||      |||
93 IAspSerGluThrIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIu 110
      |||      |||      |||      |||      |||      |||      |||
347 CAGGGAAGAGCGCTGCGCGCGACAGCTCTGCGGCTGCAAGAGAC 396
      |||      |||      |||      |||      |||      |||      |||
110 ergIAspAsp.....GluAspAspThrAspGlyAlaIuIuIuIuIu 122
      |||      |||      |||      |||      |||      |||      |||
397 CCGCGCGCGCGAG.....CACTGGCGACAGCGCGCTTCACACAGCC 437
      |||      |||      |||      |||      |||      |||      |||
123 PheValSerGluAsnSerAsnAsnIuIuIuIuIuIuIuIuIuIuIu 139
      |||      |||      |||      |||      |||      |||      |||
438 CTCGAGATGAGCGCGCGGTGATGCGCGCGCGGTGATGATGATGATG 487
      |||      |||      |||      |||      |||      |||      |||
139 rGluIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIu 156
      |||      |||      |||      |||      |||      |||      |||
488 GGTCAAGTGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 537
      |||      |||      |||      |||      |||      |||      |||
156 yspPheArgCysProIaIuIuIuIuIuIuIuIuIuIuIuIuIuIuIu 172
      |||      |||      |||      |||      |||      |||      |||
538 AAGGACACACGCGCTGACG.....CGCCACAGCGCGCGCGCGCG 578
      |||      |||      |||      |||      |||      |||      |||
173 LysAsnGlyIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIu 189
      |||      |||      |||      |||      |||      |||      |||
579 CAGGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 628
      |||      |||      |||      |||      |||      |||      |||
189 IArgAsnGlnHisIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIu 206
      |||      |||      |||      |||      |||      |||      |||
629 GCGGCAATACACCTGCGCGCGGTGCGAGCGCGCGCGCGCGCGCGCG 678
      |||      |||      |||      |||      |||      |||      |||
206 ysgIAsnIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIu 222
      |||      |||      |||      |||      |||      |||      |||
679 ACCTACAGAGGTGATGATGATGATGATGATGATGATGATGATGATG 728
      |||      |||      |||      |||      |||      |||      |||
223 ThrTyriHisIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIu 239
      |||      |||      |||      |||      |||      |||      |||
729 AGGACGCGCGCGGTGAACAGAGAGTGAAGTGAAGTGAAGTGAAGTGA 778
      |||      |||      |||      |||      |||      |||      |||
239 nAlaGlyIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIuIu 256
      |||      |||      |||      |||      |||      |||      |||
779 TCCAGTGCAGAGTGGCGAGCGAGCTGAAGCGGTGATCCAGTGGCTGA 828
      |||      |||      |||      |||      |||      |||      |||

```

```

256 hevalCysLysValLysSerAspAlaGlnProHisIleGlnTrpIleLys 272
829 GCGGTGGAG.....TACGGCGCCAGAGCGCCCAACATC 863
273 HisValGlnLysAsnGlySerLysTrpGlyProAspGly..... 285
864 CACCATGATGTGGGGCGGCGAAGTTGTGTGTCTGCTGCCACGGGTGACG 913
286 .....LeuProTyrLeuLysV 291
914 TG.....TGTGCGCGCCCGACGGCTCTACCTCATTAAGCTGCTCATC 957
221 allLeuLysHisSerGlyLysAsnSerSerAsnAlaGluValLeuAlaLeu 307
958 ACCGTGCGCCCGACGAGCATGGGCGCATGATCATCTGCTTGGCGCCAA 1007
308 PheAsnValThrGlnAlaAspAlaGlyLysTrpLysValSerAs 324
1008 CACCATGGGCTACAGCTTCGGCGCGCTCTCTACCGCTGCGCCAGACC 1057
324 hTyrIleGlyGlnAlaAsnGlnSerAlaTrpLeuThrValLeuProLysG 341
1058 CAAACCGCCGAGG 1071
341 LnglnAlaProGly 345

seq.name: /cgn2.6/prodata/2/1aa/5a_COMB.pep:US-07-640-029-1

seq_documentation_block:
: Sequence 1, Application US/07640029
: Patent No. 5229501
: GENERAL INFORMATION:
: APPLICANT: Kiefer, Michael C.
: APPLICANT: Valenzuela, Pablo D.T.
: APPLICANT: Barr, Philip J.
: TITLE OF INVENTION: Expression and Use of Human Fibroblast
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Chilton Corporation
: STREET: 4560 Horton Street
: CITY: Emeryville
: STATE: California
: COUNTRY: USA
: ZIP: 94608
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/640,029
: FILING DATE: 19910111
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: McClung, Barbara G.
: REGISTRATION NUMBER: 33,113
: REFERENCE/DOCKET NUMBER: CH-165
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 510-601-2708
: TELEFAX: 510-655-3542
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 816 amino acids
: TYPE: AMINO ACID
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-07-640-029-1

alignment_scores:
Quality: 431.50 Length: 324

```

```

Ratio: 2.007 Gaps: 9
Percent Similarity: 66.358 Percent Identity: 31.790

alignment_block:
US-09-598-042a-2 x US-07-640-029-1 ..
Align seg 1/1 to: US-07-640-029-1 from: 1 to: 816

112 GCGCGCATATGCGGCTGCGATGTCGCCAGTGAAGGGGAGCCGCGCGCT 161
|| .....|||
48 GLysArpLeuGlnLeuArgLysArgLeuArgAspArgValGlnSerI 64
162 GACCATGTGGACCAAGATGGCGGCAACCATTCACAGCGCTGGACCGCT 211
::: || .....|||
64 eAsn...TrpLeuArgAspGlyValGlnLeuAlaGlnSer...AsnArgT 79
212 TCGCGCTGCTGCCGAGGCGGTGAAGGTGAAGCATGAAGGTGAGCGGAGAT 261
|||::: :: .....|||
79 hTyrGlnLeuThrGlyGlnGluValGlnValGlnAspSerValProAlaAsp 95
262 GCGGCGGTGTACGTGTGCAAGGCCACCAACGCGCTTGGCGAGCTGAGCGT 311
::: |||::: ||| .....|||
96 SerGlyLeuTyrAlaCysValThrSerSerProSerGlySer...AspThr 111
312 CACATACACCCCTGCTGCTGATGACATTTAGCCAGGAGGAGAGGCC 361
::: ||| || .....|||
111 ThrTyrPheSerValAsnValSerAspAlaLeuProSerSerGluAspA 128
362 TGGGCGCCGACAGCTCTGCGGGGTCAAG..... 393
|||::: |||::: |||::: |||
128 sPAspAspAspAspAspAspSerSerSerGlnGlnLysGlnLysGlnTrp 144
394 GACCGCGCCGACGACAGTGGGCGACGCGCGCTTCACAGCCCTCCAA 443
|| .....::: :: || .....|||
145 AspAsnThrLysProAsnProValAlaProTyrThrSerProGlnLys 161
444 GATGAGCGCGCGGTGATGCGACGCGCGCGGTGATGCGCGCGGTGAT 493
|||::: ||| .....|||
161 sMetGlnLysLysLeuHisAlaValProAlaAlaLysThrValLysPheL 178
494 AGTGGGTGGCCAGCGGCGACCGCTGCGCGCGCGCATGATGAAGAGC 543
|||::: ||| .....|||
178 yScysProSerSerGlyThrProAsnProThrLeuArgTrpLeuLysAsn 194
544 GACGAGCGCTTGACGCGCGCCAG.....GCGGTGAGCGCCAG 581
::: :: .....|||
195 GlnLysGlnPhe...LysProAspHisArgTyrGlnTyrLysValArg 210
582 GAAGAAGAAGTGACATGACGCTGAAGAAGCTGGCGCGCGGAGAGCAGG 631
| .....||| .....|||
210 gTyrAlaThrTrpSerIleIleMetAspSerValValProSerAspArg 227
632 GCAATATACCTGCGCGGTGTCGAACCGCGCGCGCGCGCATCAACGCGAC 681
|||::: ||| .....|||
227 LysAsnTyrThrCysLysLeuAlaGlnAsnGlnTyrGlySerLysLeuHis 243
682 TACAAGGTGATGTGATCCAGCGGACCGCTTCCAAAGCGCTGTCACAGG 731
|||::: ||| .....|||
244 TyrGlnLeuAspValValGlnArgSerProHisArgProIleGlnLys 260
732 CACGACCGCGGTGAACACGAGGTGACTGGGGGAGACACAGCTCTTCC 781
: .....||| .....|||
260 aGlyLeuProAlaAsnLysThrValAlaLeuLysSerAsnValGlnPheM 277
782 AGTGAAGGTGGCGACGACAGTGAAGCGCGGTGATGCGGTGAAGCGC 831
::: ||| .....|||
277 eCysLysValLysSerAspProGlnProHisIleGlnTrpLeuLysHis 293
832 GTGAGGTACGGCGCGGAGGCGCGCAAACTCCACCATGATGTGGCGCG 881
::: ||| .....|||
294 IleGlnTrpGly.....SerLysIleGlyProAspAs 304
882 CCAGAAAGTTTGTG...GTGCTGCCACGCGGTGACGTTGTGGCGCGCG 928

```

```

304 nleuprotlyrValInleuLysThrIaGlyValasnThrIrasPL 321
929 ACGGCTCTACATCAATAGCTGCTCATACCCGCGCCGCCAGAGAGAT 978
321 ysglmetpheVal.....LeuHisLeuArgasnValSerPhegluAsp 335
979 GCGGCGATGATCATCTGCTTGCGGCCAACAACATGGGCTACAGCTTCCG 1028
336 llaGlyIuTrThrCysLeuAlaGlyasnSerIleGlyLeuSerHisH1 352
1029 CAGCGCCTCTCTCACCGTGTG 1050
352 sSerIaItrPheThrValLeu 359

seq_name: /cgn2_6/prodata/2/laa/5A_COMB.pep:US-07-921-807B-3

seq_documentation_block:
; Sequence 3, Application US/07921807B
; Patent No. 5474914
; GENERAL INFORMATION:
; APPLICANT: SPAETE, RICHARD
; TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESS: CHIRON CORPORATION
; STREET: 4560 Horton Street - R440
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/921,807B
; FILING DATE: 29-SEP-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCLUNG, BARBARA G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0209.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 820 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-921-807B-3

alignment_scores:
      Quality: 427.50      Length: 440
      Ratio: 1.595      Gaps: 14
Percent Similarity: 60.909      Percent Identity: 27.955

alignment_block:
US-09-598-042a-2 x US-07-921-807B-3 ..
Align seg 1/1 to: US-07-921-807B-3 from: 1 to: 820

112 GCGCGCACTATGCGGCTGCACTGCCAGTGGAGGGAGCCGCGCGCT 161
111 : : : : : : : : : : : : : : : : : : : : : : : : : :
48 GlYAspLeuGlnLeuArgCysArgLeuArgAspAspValGlnSerI1 64
162 GACCATGTGACCAAGATGGCGGCACATCCACAGCGGCTGGAGCGCT 211
```

```

64 eAsn...TrpLeuArgAspGlyValGlnLeuAlaGluSer...AsnArgT 79
212 TCCCGCTGTCGCCCGCAGCGGCTGAAGTGAAGCAGGTGGAGCGGAGAT 261
79 hrArgIleThrGlyGluGluValGlnValAspSerValProIaAsp 95
262 GCGCGCTGTATACGTGTGCAAGGCCACCAAGCGGCTTCGAGCGTGAAGC 311
96 SerGlyLeuTrpAlaCysValThrSerProSerIleGlySer...AspTh 111
312 CAACTAACCCCTGCTGCTGTGATGACATTAGCCAGGAAGAGAGACC 361
111 rThrTrPheSerValasnValSerAspAlaLeuProSerSerGluAsp 128
362 TGGGCGCCGACAGCTCCTGCGGCTCAAGAG.....GACCCGCC 402
128 spAspAspAspAspSerSerSerGluGluLysGluThrAspAsnThr 144
403 AGCCAGCAGTGGCGACGACCGCGCTTCACACAGCCCTCAAGATGAGCG 452
145 LysProasnProValAlaProTyrTrpThrSerProGluLysMetGly 161
453 CCGGATATGCGACGCGCCGCTGGTACCTCGCGGCTCAAGTGCCTGG 502
161 sLyLeuHisAlaValProAlaAlaLysThrValLysPheLysCysPro 178
503 CAGCGGCGACCCCTCGCGCCGACATCACTGAGTGAAGAGACAGCCAGCC 552
178 erSerGlyThrProasnProThrLeuArgTrpLeuLysasnGlyLysGlu 194
553 TTGACGCGCCACAG.....GCCGCTGAGCCAGGAAGAA 590
195 Phe...LysProasnProHisArgIleGlyGlyLysValArgTyrAlaTh 210
591 GTGACACCTAGCCTGAAGAAGCTGCGCGCGAGAGACGCGCAATATCA 640
210 rTrpSerIleIleMetAspSerValValProSerAspLysGlyAsnTyr 227
641 CGTGCGCGCTGTGACCGCGCGCGCGCATCAAGCCAGCTCAAGCTG 690
227 hrcysIleValGluasnGluTrpIleYserIleasnHisThrTyrGlnLeu 243
691 GATGTGATCCAGCGACCGCTCCAGCCCGCTGCTCACAGCAGCAGCC 740
244 AspValValGlnArgSerProHisArgProIleLeuGlnAlaGlyLeuPr 260
741 CGTGAACACAGCGTGTGACTTGGGGGAGACAGCTCTTCCAGTGCAGG 790
260 oAlaAsnLysThrValAlaLeuGlySerAsnValGluPheMetCysLysV 277
791 TGGCAGCAGCATGAAGCGGCTGATCCAGTGGCTGAAGCGCTGAGTAC 840
277 alYrSerAspProGlnProHisIleGlnTrpLeuLysHisIleGluVal 293
841 GCGCGCAGGCGCGCAACAATCCACCATGATGTGGCGCGCAGAAGTT 890
294 .....AsnGlySerLysIleGlyProAspAsnLeuProTy 305
891 TGTG...GTGCTGCCACGCGGTGACGTGTGCTGCGCGCCGACGCGCT 937
305 rValGlnIleLeuLysThrAlaGlyVal.....AspThrThAspLysG 320
938 ACCTCAATTAAGTGTCTCATCACCCTGCGCCGCCAGCAGATGCGGCGATG 987
320 lmetGluValIleuHisLeuArgasnValSerPheGluAspAlaGlyGlu 336
988 TACATCTGCTTGGCGCAACACCATGGGCTTACAGCTTCCGACAGCGCTT 1037
337 TyrThrCysLeuAlaGlyasnSerIleGlyLeuSerHisHisSerIaTr 353
1038 CMTACCGTGTGCA.....GACCCAAACGCGCAGGCGCACCTGTGG 1081
```



```

841 GGGCGGCGAGGCGCCAGCAACTCCACCATGATGTGGCGGCGGAGGATT 890
294 .....AsnGlySerLysIleGlyProAspSerLeuProTy 305
891 TGTG...GTGTGCGCCAGGGGTGACGTGTGTGCGCGCGCGAGCGCTCCT 937
305 rValGlnIleLeuLysThrAlaGlyVal.....AsnThrThraspysg 320
938 ACCGCAATAGCTGCTCATCACCGCTGCCCGCCGAGAGATGGCGGCGATG 987
320 lMetGluValIleuHisLeuAlaGlnValSerPheGluAspAlaGlyGlu 336
988 TACATGTGCTGTGGCGCAACACCATGAGCTTACAGCTTCGCGACGCGCTT 1037
337 TTTThrcysLeuAlaGlyAsnSerIleGlyLeuSerHisSerAlaTr 353
1038 CMTACCGTGTGTGCA.....GACCCAAAGCCGCGGCGGCGACCTGTGG 1081
353 PleuThrValLeuGluAlaLeuGluGluArgPro..... 364
1082 CTTCTGTGCTGCGGCGCCATGACCTGCGCGCGCTGTGTCATCGGCATC 1131
365 .....AlaValMetThrSerProLeuGlyLeuGluIleIleIleTyr 378
1132 CGAGCGCGCGCTGTCTTCATC.....CTGGGCAACCTGTGCTGTG 1172
379 CysThrGlyAlaPheLeuIleSerCysMetValGlySerValIleValTy 395
1173 GCTTTGGCAGGCGCCAGAGAGCGG.....TGCA 1201
395 rLysMetLysSerGlyThrLysLysSerAspPheHisSerGluMetAlaV 412
1202 CCCCCGCGCTGCCCTCCCTGCTGCGGCGACCGCGCGCGGAGCGGCC 1251
412 alHisLysLeuAlaLysSerIleProLeuArgArgGluValThrValSer 428
1252 CGGACCGCGCGGAGAGAGAGAC.....CTCCCTC 1283
429 AlaAspSerThrAlaSerMetAsnSerGlyValLeuLeuValArgPro 445
1284 GTTGGCGCGCGCTCAGCGCTGCGCTGTGTGGGCGCTGTGAGAGAGATG 1333
445 rArgLeuSerSerSerGlyThrProMetLeuAlaGlyValSerGluTyrG 462
1334 GGTCTCCGCGACCGCCCGAG 1353
462 lLeuProGluAspProArg 468
seq_name: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:US-08-439-992A-1
seq_documentation_block:
; Sequence 1, Application US/08439992A
; Patent No. 6255454
; GENERAL INFORMATION:
; APPLICANT: Kieffer, Michael C.
; APPLICANT: Pablo, Valenzuela D.T.
; TITLE OF INVENTION: Expression and Use of Human Fibroblast
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/439,992A
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Chung, Ling-Fong
; REGISTRATION NUMBER: 36,482
; REFERENCE/DOCKET NUMBER: 0165,004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-923-2704
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 820 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-439-992A-1

alignment_scores:
Quality: 427.50 Length: 440
Ratio: 1.595 Gaps: 14
Percent Similarity: 60.909 Percent Identity: 27.955

alignment_block:
US-09-598-042A-2 x US-08-439-992A-1 ..
Align seg 1/1 to: US-08-439-992A-1 from: 1 to: 820

112 GGGCGGCAATGCGGCTGCGAGTCCAGTGAGGAGGAGCGCGCGCT 161
48 GlyAspLeuLeuGluLeuArgCysArgLeuValArgAspValGlnSerI 64
162 GACCAATGTGACCAAGATGCGCGGCGACATCCACAGCGGTGAGCGCT 211
64 eAsn...TriPLeuArgAspGlyValGlnLeuAlaGlnSer...AsnArgT 79
212 TCCGCGTGTGCGCGCGAGGCGTGAAGTGAAGCAGGTGAGCGGAGAT 261
79 hrArgIleThrGlyGluGluValGlnValGlnAspSerValProAlaAsp 95
262 GCGCGGTGTGACTGTGCAAGCGCCACCAACGCTTGCGCAGCTGAGCGT 311
96 SerGlyLeuThrValAcysValThrSerSerProSerGlySer...AspTh 111
312 CAACCTACACCCCTGCTGCTGATGACATTACCGCAGGAGAGAGAGCC 361
111 rTrpThrPheSerValAsnValSerAspAlaLeuProSerSerGluAspA 128
362 TGGGCGCGCGACAGCTCTCTGCGGCGTCAAGAG.....GACCCCGCC 402
128 spAspAspAspAspAspSerSerSerGluGluLysGluThrAspAsnThr 144
403 AGCCACAGTGGGCGACGCGCGCTTCACACAGACCCCTCCACAGATGAGCG 452
145 LysProAsnProValAlaProTyrTrpThrSerProGluLysMetIuTy 161
453 CCGGCTGATGCGACGCGCGCGTGGGTAGCTGCGCGGCTCAATGGGTGG 502
161 sLysLeuHisAlaValProAlaAlaLysThrValLysPheLysCysProS 178
503 CGAGCGGCGACCGCTGCGCGCGATCAGCTGATGAAGAGAGAGAGAGCC 552
178 erserGlyThrProAsnProThrLeuArgTrpLeuLysAsnGlyGlu 194
553 TTGACGCGCGCCAGAG.....GCCGCTGAGCCCAAGAGAGAGAA 590
195 Phe...LysProAspHisArgIleGlyGlyTyrLysValArgTyrAlaTh 210
591 GTGGACATGAGCGCTGAAGAGACCTGCGCGCGAGGAGAGGAGCAATPAC 640
210 rTrpSerIleIleMetAspSerValValProSerAspLysGlyAsnTyrT 227

```

```

641 CCGCGCGCTGTGAACCGCGCGGCGCATCATACGACGACATTAACAGGNG 690
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
227 hCysIleValIgluAsnGluTyrGlySerIleAsnHisThrTyrGlnLeu 243
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
691 GATGTATCAGCGGACCGGCTTCAAGCCGCTGTCAAGGACGACGACCC 740
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
244 AspValIgluArgSerProHisArgProIleLeuGlnIleGlyLeuArg 260
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
741 CGTGAACAGCGAGTGGAGCTTGGGGGAGCAAGCTCTTCCAGTGAAG 790
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
260 AlaAsnLysThrValIleAlaLeuGlySerAsnValGluPheMetCysLys 277
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
791 TCGGAGCAGCGAGCGGCTGATCCAGTGGCTGAAGCGCGGAGGTAC 840
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
277 ALTYSerSproGlnProHisIleGlnTrpLeuLysHisIleGluVal 293
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
841 GCGCGCGAGCGCGCCCACTCCACATCATGATGTGGCGCGCAGAAGTT 890
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
294 .....AsnLysSerLysIleGlyProAspAsnLeuProty 305
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
891 TGTG...GTGTCGCCAGCGGTGACGTGTGTCGCGCGCCGACGCTCT 937
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
305 rValGlnIleLeuLysThrAlaGlyVal.....AsnThrThrAspLysG 320
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
938 ACCTCAATAGCTGTATCACCCTGCGCGCGCGGACGACGATCGGGCATG 987
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
320 LmetGluValLeuHisLeuAlaArgAsnValSerPheGlnAspAlaGlyL 336
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
988 TACATTCCTGCTTGGCGCAACACATGAGCTTACGCTCCGACGCGCTT 1037
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
337 TyrThrCysLeuAlaGlyAsnSerIleGlyLeuSerHisHisSerLys 353
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1038 CCTCACCGCTGTGCA.....GACCCAAACCGCGGACGACCTGTGG 1081
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
353 pLeuThrValLeuGluAlaLeuGlnLysLysPro..... 364
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1082 CCTCCCTGCTGCGGCGCATACCTGCGCGCGCTGCTCATCGCATC 1131
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
365 .....AlaValMetThrSerProLeuTyrLeuGlnIleIleTyr 378
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1132 CCAGCGCGCGCTGTCTTATC.....CTGGGACCGCTGCTCTGTG 1172
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
379 CysThrGlyAlaPheLeuIleSerCysMetValGlySerValIleVal 395
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1173 GCTTTCGCGCGCGCAGAGAGCG.....TGCA 1201
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
395 rLysMetLysSerGlyThrLysLysSerAspPheHisSerGlnMetAla 412
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1202 CCGCGCGCGCTGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCG 1251
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
412 ALHisLysLeuAlaLysSerIleProLeuAlaArgGlnValThrValSer 428
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1252 CCGGACCGCGAGCGAGACAGGAC.....CTGCCCTC 1283
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
429 AlaAspSerSerAlaSerMetLysSerGlyValLeuLeuValAlaArg 445
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1284 GTTGGCGCGCTGAGCGCTGCGCTGCGCGCGCTGTGAGGACAGCAT 1333
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
445 rArgLeuSerSerSerGlyThrProMetLeuAlaGlyValSerGluTyr 462
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1334 GGTCTCCGCGACCGCGCGCGCG 1353
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
462 LuleuProGluAspProArg 468
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
seq_name: /cgn2_6/ptodata/2/1aa/PCrus_COMB.pep:PCT-US93-05703-2
seq_documentation_block:
: Sequence 2, Application PC/TUS9305703
: GENERAL INFORMATION:
: APPLICANT: Nova, Michael P.
: APPLICANT: Gonzalez, Ana-Maria
: APPLICANT: Baird, J. Andrew

```

```

: TITLE OF INVENTION: Process for Detection of Neoplastic
: TITLE OF INVENTION: Disease
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fitch, Even, Tabin & Flannery
: STREET: 135 South LaSalle Street, Suite 900
: CITY: Chicago
: STATE: IL
: COUNTRY: USA
: ZIP: 60603
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US93/05703
: FILING DATE: 19930614
: CLASSIFICATION:
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 07/900,646
: FILING DATE: 18-JUN-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Seidman, Stephanie
: REGISTRATION NUMBER: 33,779
: REFERENCE/DOCKET NUMBER: 54625PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-552-1311
: TELEFAX: 619-552-0095
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 351 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: FEATURE:
: NAME/KEY: Protein
: LOCATION: 1..351
: OTHER INFORMATION:
: OTHER INFORMATION: extracellular domain of an FGFR1"
: PCT-US93-05703-2

alignment_scores:
Quality: 425.50 Length: 321
Ratio: 1.970 Gaps: 9
Percent Similarity: 67.290 Percent Identity: 31.776

alignment_block:
US-09-598-042a-2 x PCT-US93-05703-2 ..
Align seg 1/1 to: PCT-US93-05703-2 from: 1 to: 351

112 GCGCGCATATGCGGCTGCGAGTCCCGAGTGGAGGGGACCGCGCGCT 161
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 GlyAspLeuLeuGlnLeuArgCysArgLeuArgAspValGlnSerIle 41
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
162 GAGCAGTGTGACCAAGATGCGCGCACCATCCACAGCGGCTGAGCGCGT 211
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
41 easn...TripleuArgAspGlyAlaGlnLeuAlaGluSer...AsnArgT 56
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
212 TCCGCGTGTGCGCGAGGGGCTGAGGTGAAGCAGGTGAGCGGAGAGT 261
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
56 hrArgLethrGlyGluGluValGlnValGlnAspSerValProAlaAsp 72
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
262 GCGCGCGTGTAGCTGTGCAAGGCGACCAACGCTTGGGAGCGCTGAGC 311
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73 SerGlyLeuTyrAlaCysValThrSerSerProSerGlySer...AspTh 88
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
312 CAAGTACACCTGCTGCTGTGATGATCAATTAGCCGAGGAGAGAGAGCC 361
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
88 rThrTyrPheSerValAsnValSerAspAlaLeuProSerSerGluAspA 105

```





```

161 tglutylsleuHsAlaValProAlaAlaLysThrValLysPheLysC 178
497 GCGTGGCCAGCGGGACCTCGGCCGACATCAGCTGGATGAAGACGAC 546
178 ysPserSerSerGlyThrProAsnProThrLeuAlaGlyThrLeuLysAsnGly 194
547 CAGGCGCTTGCAGCGGGCCAGAG.....GCGGCTGAGCCCGAGAA 584
195 LysGluPhe...LysProAsnHisArgIleGlyGlyTyrLysValArgTyr 210
585 GAAGAAGTGGACACTGAGCCTGAGAACTGGCGCGGAGACGCGGAGA 634
210 rAlaThrTrpSerIleIleMetAspSerValAlaProSerAspLysGlyA 227
635 AATACACCTGGCGCGCTGTCGAAACCGCGGCGCGCATCAACGCGACCTAC 684
227 sntYThrTyrSileValGluAsnGluTyrGlySerIleAsnHisThrTyr 243
685 AAGGTGGATGTATCGACGCGGACCGCTTCCAGGCGCGCTGTACACAGCAC 734
244 GlnLeuAspValAlaGluLysSerProHisArgProIleLeuGlnAlaG 260
735 GCACCCCGTGAACACAGCGGTGACTTCGGGGGAGCACGTCCTTCAGT 784
260 yLeuProAlaAsnLysThrValAlaLeuGlySerAsnValGluPheMetC 277
785 GCAGGTGGCGACGACAGTGAAGCCGCGTATCCAGTGGCTGAAGCGCGGTG 834
277 yLeuValLysSerAspProGlnProHisIleGlnIlePheLysHisIle 293
835 GAGTACGGGGCGGAGGCGCGCACACTCCACATGATGTGGCGCGGCA 884
294 GluVal.....AsnGlySerLysIleGlyProAspAsn 305
885 GAAGTTTGTG...GTGCTGCCACGCGGTGACGTGTGTGGCGCGCGGACG 931
305 uproGlyrValGlnIleLeuLysThrAlaGlyAla.....AsnThrThra 320
932 GCTCCTACCTCAATTAAGCTGTGCATACCGCTGCCGCGACGACGATGGC 981
320 sPlysGluMetGluValLeuHisIleuAlaGAsnValSerPheGluAspAla 336
982 GGCATGTACATCTGCTGGCGGACACACCATGGGCTACAGCTTCGCGAG 1031
337 GluGluTyrTyrThrCysLeuAlaGlyAsnSerIleGlyLeuSerHisS 353
1032 CGCCTTCTCTACCGCTGTGTGCA.....GACCCAAACCGCGGAGCGCAC 1075
353 rAlaTrpLeuThrValLeuGlnAlaLeuGluGluArgPro..... 366
1076 CTGTGGCTCTCTGCTGCGGACATGACCTGCGCGCGCTGTGCATC 1125
367 .....AlaValMetThrSerProLeuTyrLeuGluIleIle 378
1126 GGCATCCGACGCGGCGCTGTCTTCATC.....CTGGACCGCTGTGCT 1166
379 lIeTyrCysThrGlyAlaPheLeuIleSerCysMetValGlySerValI 395
1167 CCTGTGGCTTTTGCAGCGCCAGAAAGAGCG..... 1197
395 eValTyrLysMetLysSerGlyThrLysLysSerAspPheHisSerGlnM 412
1198 ..TGACACCCCGCGCTGCGCTGCGCTGGGACGCGCGCGCGG 1245
412 eValAlaHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 428
1246 AGCGCGCGGAGCGGACGCGGAGACAGAGAC.....CT 1277
429 ValSerAlaAspSerSerAlaSerMetAspSerGlyValLeuLeuValAr 445
1278 TCCTCGTTTGCGCGCTCAGCGCTGCGCTGCTGGGCTGTGAGG 1327
445 gProSerArgLeuSerSerSerGlyThrProMetLeuAlaGlyValSerC 462

```

```

1328 AGCATGGGTCTCCGGCAGCCCCCAG 1353
462 lutyGluLeuProGluAspProArg 470
seq_name: /cgn2_6/plodata/2/laa/5A_COMB.pep:US-07-921-807B-4
seq_documentation_block:
; Sequence 4, Application US/07921807B
; Patent No. 5474914
; GENERAL INFORMATION:
; APPLICANT: SPAETE, RICHARD
; TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
; TITLE OF INVENTION: OF VIRAL PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: 4560 Horton Street - R440
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/921,807B
; FILING DATE: 29-SEP-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCLOUD, BARBARA G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0209, 001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 822 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-921-807B-4
alignment_scores:
Quality: 423.50 Length: 442
Ratio: 1.580 Gaps: 14
Percent Similarity: 60.633 Percent Identity: 27.602
alignment_block:
US-09-598-042A-2 x US-07-921-807B-4 ..
Align seg 1/1 to: US-07-921-807B-4 from: 1 to: 822
112 GCGCGCATATGCGGCTGCAGTGCACATGAGAGGAGGAGCCGCGCGCGCT 161
111 ..... 111
48 GluAspLeuLeuGlnLeuArgCysArgLeuArgAspAspValGlnSerI 64
162 GACCATGTGACCAAGATGCGCGCACATCCACACGCGCTGAGCGCT 211
64 easn...TriLeuArgAspGlyValGlnLeuAlaGlnSer...AsnArgT 79
212 TCCGCGCTGCTGCGCGGAGGCTGAAGTGAAGCAGGTGAGCGGAGGAT 261
79 hrrArgIleThrGlyGluGluValGlnValGlnAspSerValProAlaAsp 95
262 GCGCGGTGATGCTGTGCAAGGCGCACACGAGCGCTTGGAGCGCTGAGCG 311
96 SerGlyLeuTyrAlaCysValThrSerSerProSerGlySer...AspTh 111

```

```

312 CAACCTACACCTCTGCTGCTGATGACATTACGCCAGAGAGAGGCC 361
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
111 rThTyPheSerValAsnValSerAspAlaLeuProSerSerLeuAsp 128
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
362 TGGGGCCCGCAGAGCTCTGGGGGATGAGAGAGCCCGCC..... 402
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
128 spAspAspAspAspSerSerSerSerLeuGluGluSerGluThrAsp 144
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
403 .....AGCAGACAGTGGGACGACGAGCCGCTTCACACAGCCCTCCAA 446
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
145 LysProAsnArgMetProValAlaProTyrThrSerProGluLysMe 161
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
447 GAGCGCGCGGGTATCGACGCGCGCGGGGATGCTCGGCGGGTCAAGT 496
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
161 fGluLysLysLeuHisAlaValProAlaAlaLysThrValLysPheLys 178
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
497 GCGTGGCAGCGGCGACCTCGCGCGCGACATCAGCTGATGATGAGAGAC 546
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
178 YsProSerSerGlyThrProAsnProThrLeuArgThrPheLysAsnGly 194
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
547 GAGCGCTTGACGCGCCAGAG.....GCCGCTGAGCCCGAGAA 584
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
195 LysGluPhe...LysProAspHisArgIleGlyGlyTyrLysValArgTyr 210
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
585 GAAGAAGTGGACACTGAGCCGAGAGAAGCTCGCGCGGAGAGACGCGGCA 634
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
210 rAlThrTrpSerIleIleMetAspSerValValProSerAspLysGly 227
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
635 AATACACCTCGCGCGGTGTGAAACCGCGCGGCGGACATCAACGCGCATC 684
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
227 snTyThrCysIleValGluAsnGluTyrGlySerIleAsnHisThrTyr 243
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
685 AAGTGGATGTGATCGACGCGCGCGGTTCACAGCCCGGTCTCAAGGCGAC 734
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
244 GluLeuAspValValGluArgSerProHisArgProIleLeuGlnAlaGly 260
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
735 GCACCGCGCTGAAACGACGAGCTGAGTGGGGGAGACGACGCTTCCAGT 784
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
260 YLeuProAlaAsnLysThrValAlaLeuGlySerAsnValAlaGluPheMetC 277
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
785 GCAAGGTGCGCAGCAGCTGAAAGCGGATGATCAGTGGCTGAAGCGCGTG 834
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
277 YsLysValTyrSerAspProGlnProHisIleGlnTrpLeuLysHisIle 293
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
835 GAGTACGCGCGCGGAGCGCGCGCACACATCCACATCATGTGGCGCGGCA 884
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
294 GluVal.....AsnGlySerLysIleGlyProAspAspLe 305
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
885 GAAGTTTGTG...GTGTCGCCCGGAGGAGCTGTGTCGGCGCGCGCAG 931
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
305 uProTyrValGlnIleLeuLysThrAlaGlyVal.....AsnThrTrp 320
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
932 GCTCCTACCTCAATAGCTGTATCAGCCGCTGCCCGCGCAGAGAGATGCG 981
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
320 sPlyGluMetGluValLeuHisLeuArgAsnValSerPheGluAspAla 336
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
982 GGCATGTACATCTGCTGGCGCGCACACCATGGCTACACTTCGCGAG 1031
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
337 GluGluTyrThrCysLeuAlaGlyAsnSerIleGlyLeuSerHisHisSe 353
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1032 GCGCTTCCCTCAACGCTGCGCA.....GACCAAAACCGCAGGGCGCAC 1075
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
353 rAlaThrPheThrValLeuGluAlaLeuGluGlnArgPro..... 366
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1076 CTGGGCGCTCTGCTGCTGCGGCACTACCTGCGCGCGCGCTGTATC 1125
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
367 .....AlaValMetThrSerProLeuTyrLeuGlnIleLeu 378
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1126 GGCATCCAGCGCGCGCTGTCTTCAATC.....CTGGGCGACCGCTCT 1166
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
379 IleTyrCysThrGlyAlaPheLeuIleSerCysMetValGlySerValI 395

```

```

1167 CTGTGCTTTGGCAGCGCCAGAGAGCGG..... 1197
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
395 eValTyrMetLysSerGlyThrLysLysSerAspPheHisSerGln 412
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1198 ..TGACCCCCCGCGCTGCCCTCCCTGCTGGCGACCGCCCGCGGG 1245
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
412 eValValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 428
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1246 ACGGCGCGCGCAGCGCGAGGAGACAGAGAC.....CT 1277
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
429 ValSerAlaAspSerSerAlaSerMetAsnSerGlyValLeuValArg 445
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1278 TCCCTGTGGCGCGCGCTGACGCGCTGCGCTGTGGGGCTGTGAGG 1327
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
445 gProSerArgLysSerSerSerGlyThrProMetLeuAlaGlyValSerG 462
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1328 AGCATGGGCTCCGCGCAGCGCCCGCAG 1353
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
462 LutyGluLeuProGluAspProArg 470

```

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-08-459-296-2

```

seq_documentation_block:
; Sequence 2, Application US/08459296
; Patent No. 5670323
; GENERAL INFORMATION:
; APPLICANT: No. 5670323a, Michael
; APPLICANT: Gonzalez, Ana-Maria
; APPLICANT: Baird, Andrew
; TITLE OF INVENTION: PROCESS FOR DETECTION OF NEOPLASTIC
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,296
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mastrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7573-024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 822 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-459-296-2

```

```

alignment_scores:
  Quality: 423.50      Length: 442
  Ratio: 1.580         Gaps: 14
  Percent Similarity: 60.633      Percent Identity: 27.602

```

alignment\_block:  
US-09-598-042a-2 x US-08-459-296-2 ..

Align seg 1/1 to: US-08-459-296-2 from: 1 to: 822

```

112 GCGCCGACATATGCGGCTGAGTGCAGTGCAGTGCAGGAGGAGCCGCCGCT 161
    ||| :|||:||||| :||| :|||
48 GLYAspLeuLeuGlnLeuArgCysArgLeuArgAspAspValGlnSer11 64
162 GACGATGTGGACCAAGAGATGGCCGACCATCCACAGCGGCTGGAGCCGT 211
    :||| :|||:||||| :||| :|||
64 easn...TrpLeuArgAspGlyValGlnLeuAlaGlnSer...AsnArgT 79
212 TCCGCGCTGCTGCGGAGGCTGAAGTGAAGTGAAGTGAAGGAGGAGAT 261
    |||:||||| :||| :|||:||||| :|||
79 hrArgIleThrGlyGlnGluValGlnValGlnAspSerValProAlaAsp 95
262 GCGCGGCTGTACGTGTGCAAGGCCACCAAGCGCTTGCAGCGCTGAGCGT 311
    :|||:||||| :||| :|||:||||| :|||
96 SerGlyLeuThrValGlnValThrSerSerProSerGlySer...AspThr 111
312 CAACATACACCCCTGCTGCTGATGACATTAAGCCAGGAGGAAGAGAGCC 361
    :|||:||||| :||| :|||:||||| :|||
111 rThrArgPheSerValAsnValSerAspAlaLeuProSerSerGlnAspA 128
362 TGGGCGCCGACAGCTCTGCTGGGGTCAAGAGAGCCCGCC... 402
    |||:||||| :||| :|||:||||| :|||
128 sPAspAspAspAspAspSerSerSerGlnGlnGlnGlnThrAspAsnThr 144
403 .....AGCCAGACGTGGGACAGCCGCGCTTACACAGCCCTCCAGAT 446
    :|||:||||| :||| :|||:||||| :|||
145 LysProAsnArgMetProValAlaProTyrTrpThrSerProGlnLysMet 161
447 GAGGCGCGGGGTGATGCGACGCGCGCTGGGTAGCTCCGTCCGCGCTCAAGT 496
    :|||:||||| :||| :|||:||||| :|||
161 tGlnLysLysLeuHisAlaValProAlaAlaLysThrValLysPheLysC 178
497 GCGTGGCCAGCGGACCGCTGCGGCGCCGACATGATGATGAAGAGAGAGC 546
    |||:||||| :||| :|||:||||| :|||
178 yspProSerSerGlyThrProAsnProThrLeuArgTrpLeuLysAsnGly 194
547 CAGGCGCTTGCAGCGCCGACAG.....GCCGCTGAGCCGACGAA 584
    :|||:||||| :||| :|||:||||| :|||
195 LysGlnPhe...LysProAspHisArgIleGlyGlyTyrLysValAlaArgTyr 210
585 GAAGAAGTGGACACTGAGCCTGAAGAACCCTGCGCGCGAGGACAGCGCA 634
    :|||:||||| :||| :|||:||||| :|||
210 rAlaThrTrpSerIleIleMetLysSerValAlaProSerAspLysGlyA 227
635 AATACACCTGCGCGCTGTCGAACCGCGGCGCCGACATCAAGCCGACCTAC 684
    :|||:||||| :||| :|||:||||| :|||
227 snTyrThrCysIleValGlnAsnGlnTyrGlySerIleAsnHisThrTyr 243
765 GCACCCCGTGAACAGAGGTGAGTTCGGGGGACACAGCTCTTCCACT 784
    :|||:||||| :||| :|||:||||| :|||
243 yLeuProAlaAsnLysThrValAlaLeuGlySerAsnValGlnPheMetC 277
785 GCAAGCTGCGCGAGCGTGAAGCCGCGTGAAGTGCAGTGGCTGAAGCGCGG 834
    :|||:||||| :||| :|||:||||| :|||
277 yslValValTyrSerAspProGlnProHisIleGlnTrpLeuLysHisIle 293
835 GAGTACGCGCGCGAGGCGCCACAACTCCACCATCGATGTGGCGCGCA 884
    :|||:||||| :||| :|||:||||| :|||
294 GluVal.....AsnGlySerLysIleGlyProAspAsnLe 305
885 GAAGTTGTG...GTGCTGCCACGCGTACGTGTGCTGCCCGCCGAGC 931
    :|||:||||| :||| :|||:||||| :|||
305 uProTyrValGlnIleLeuLysThrAlaGlyVal.....AsnThrThrA 320
932 GCTCTACTCAATAGCTGTATACACCGCGCGCGCGAGAGAGATGGC 981
    :|||:||||| :||| :|||:||||| :|||

```

```

320 sPlySGlnMetClnValLeuHisLeuArgAsnValSerPheGlnAspAla 336
962 GGCATGATCATCTGCTTGGCGCCAAACACCATGAGCTTACAGTCCGAG 1031
    ||| ||| |||||:|||||:||||| ||| :|||
337 GlyGlnTyrThrCysLeuAlaGlnLysSerIleGlyLeuSerHisIleSe 353
1032 CGGCTTCTCCACCGTGTGCCA.....GACCCAAACCGCCGAGGCCAC 1075
    |||||:||||| ||||| :||| :|||
353 rAlaThrPheThrValLeuGlnAlaLeuGlnGluArgPro..... 366
1076 CTGTGGCTCTCTGCTGCTGCGCCACTAGCTGCTGCGCGCTGATC 1125
    :|||:||||| :||| :|||:||||| :|||
367 .....AlaValMetThrSerProLeuTyrLeuIleIle 378
1126 GGCATCCAGCGCGCTGTCTTTCATC.....CTGGGACCCCTGCT 1166
    :|||:||||| :||| :|||:||||| :|||
379 IleTyrCysThrGlyAlaPheLeuIleSerCysMetValGlySerVal11 395
1167 CCTGTGGCTTCCAGCGCCAGAGAGAGCGG..... 1197
    :|||:||||| :||| :|||:||||| :|||
395 eValTyrLysMetLysSerGlyThrLysLysSerAspPheHisSerGlnM 412
412 eAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 428
1246 ACGGCGCGGACCGGACGCGGACAGAGAC.....CT 1277
    :|||:||||| :||| :|||:||||| :|||
429 ValSerAlaAspSerSerAlaSerMetLysSerGlyValLeuValAla 445
1278 TCCGCTGTTGGCGCGCTTCCAGCGCGCTGCTGCTGCGCGCTGTGAGG 1327
    ||||| :|||:||||| :||| :|||:||||| :|||
445 gProSerArgLeuSerSerGlyThrProMetLeuAlaGlyValSerG 462
1328 AGCATGGGTCTCCGACGCGCCCGCAG 1353
462 LntyrGlnLeuProGlnAspProArg 470

seq_name: /cgn2_6/prodata/2/iae/5a_COMB.pep:US-08-441-944A-4
seq_documentation_block:
: Sequence 4, Application US/08441944A
: Patient No. 5767250
: GENERAL INFORMATION:
: APPLICANT: SPAETE, RICHARD
: TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CHIRON CORPORATION
: STREET: 4560 Horton Street - R440
: CITY: Emeryville
: STATE: CA
: COUNTRY: USA
: ZIP: 94608-2916
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/441,944A
: FILING DATE: 16-MAY-1995
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/921,807
: FILING DATE: 29-SEP-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: MCCLUNG, BARBARA G.
: REGISTRATION NUMBER: 33,113
: REFERENCE/DOCKET NUMBER: 0209.001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (510) 601-2708

```

TELEFAX: (510) 655-3542  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 822 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-441-944A-4

Alignment scores:  
 Quality: 423.50 Length: 442  
 Ratio: 1.580 Gaps: 14  
 Percent Similarity: 60.633 Percent Identity: 27.602

Alignment block:  
 US-09-598-042A-2 x US-08-441-944A-4

Align seg 1/1 to: US-08-441-944A-4 from: 1 to: 822

```

112 GCGCCCATATGCGCTGCTGAGTCCAGTGAAGGGGACCGCGCGCT 161
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
48 GlyAspLeuLeuGlnLeuArgCysArgLeuArgAspAspValGlnSer11 64
162 GACCATGTGACCAAGATGGCGCCGACCATCCACAGCGGCTGGAGCGCT 211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
64 easn...TrrLeuArgAspGlyValGlnLeuAlaGlnSer...AsnArg 79
212 TCCGCGTCTGCGCGAGGCGCTGAAGTGAAGCAAGTGAAGCGGAGGAT 261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
79 hrArgIleThrGlyGlnGluValGlnValGlnAspSerValProAlaAsp 95
262 GCGCGCGTGTACGTGTGCAAGCCACCAAGCGCTTGGCGCGTGAAGCGT 311
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
96 SerGlyLeuTyrrAlaCysValThrSerSerProSerGlySer...AspTh 111
312 CAATACACCTCTGCTGCTGATGACATTAAGCCAGGAGGAGAGAGACC 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
111 rThrTyrrPheSerValAsnValSerAspAlaLeuProSerSerGlnAsp 128
362 TGGGGCCCGACAGCTCTGCGGGGTCAAGAGACCGCGC..... 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
128 spaAspAspAspAspSerSerSerGlnGlnGlnGlnGlnGlnGlnGln 144
403 .....AGCCAGCAGTGGGCGACGCGCGCTTCACAGCCCTCCAGAT 446
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
145 LysProAsnArgMetProValAlaProTyrTrpThrSerProGlnLys 161
447 GAGCGCGCGGTATCGCACGGCGCTGGGATGCTCGTGGCGCTCAAGT 496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
161 tGlnLysLysLeuHisAlaValAlaProAlaAlaLysThrValLysPhe 178
497 GCGTGGCGACGGGCGACCTTGGCGCGACATCAAGTGAAGAGAGAC 546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
178 yspProSerSerGlyThrProAsnProThrLeuArgTrpLeuLysAsnGly 194
547 CAGCGCTTGACGGCGCCAGAG.....GCGCGTACGCCCGAGAA 584
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
195 LysGlnPhe...LysProAsnArgMetProValAlaProTyrTrpThr 210
585 GAAGAACTGACACTGAGCTGAAGAACTGCGCGGAGAGAGCGGCA 634
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
210 rAlaThrTrpSerIleIleMetAspSerValAlaProSerAspGlyGly 227
635 AATAACACTTCCCGGTGTGCAACCGCGCGGCGGCATCAACGCCACTAC 684
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
227 snTyrrThrCysIleValGlnAsnGlnTyrGlySerIleAsnHisThrTy 243
685 AAGTGATGTATGATCAAGCGGAGCGCGCTTCAACCGCGGTGTCAAGGAC 734
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
244 GlnLeuAspValValGlnArgSerProHisArgProIleLeuGlnAlaG 260

```

```

735 GCACCCCGTGAACACAGAGGTGAGTCTGGGGGAGACCAAGCTCTTCAGT 784
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
260 yLeuProAlaAsnLysThrValAlaLeuGlySerAsnValAlaGlnPheMetC 277
785 GCAAGTCCGCGACGAGCTGAAGCCGGTATTCAGTGGCTTAAGCGCGTG 834
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
277 yLysValTyrrSerAspProGlnProHisIleGlnTrpLeuLysHisIle 293
835 GAGTACGCGCGGAGCGCGCCACACTCCACATCCATGATGGGCGGCA 884
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
294 GluVal.....AsnGlySerIleGlyProAspAsnLe 305
885 GAAGTGTGTG...GTGTCGCCAGCGGTGAGCTGTGTGCGCGCCGACG 931
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
305 uProTyrValGlnIleLeuLysThrAlaGlyVal.....AsnThrThra 320
932 GCTCTACCTCAATAGCTGTCTACACCGCTGCGCGCCGACGAGAGTGC 981
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
320 sPlysGlnMetGluValLeuHisLeuArgAsnValSerPheGlnAspAla 336
982 GGCATGTACATCTGCTTGGCGCCACACCATGGGCTACAGCTTCCGCGAG 1031
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
337 GlyLysLysThrCysLeuAlaGlyAsnSerIleGlyLeuSerHisIleSe 353
1032 CGCCTTCCTACCGCTGTGCCA.....GACCCAAACCGCCAGGGCCAC 1075
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
353 rAlaTrpLeuThrValLeuGlnAlaLeuGlnGlnArgPro..... 366
1076 CTGTGGCTCTCTGCTGCTGCGCACTAGCTGCGGTGCGCGCGGTATC 1125
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
367 .....AlaValMetThrSerProLeuTyrrLeuGlnIleIle 378
1126 GGCATCCACGCGCGCGCTGTCTTCATC.....CTGGGACCGCTGT 1166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
379 IleTyrrCysThrGlyAlaPheLeuIleSerCysMetValGlySerValII 395
1167 CCTGTGCTTGTGCCAGGCGCCAGAAAGACCG..... 1197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
395 eValTyrrLysMetLysSerGlyThrLysLysSerAspPheHisSerGln 412
1198 ..TGACCCCGCGCGCTGCGCGCTCCCTGCTGCGGACCGCGCGGGG 1245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
412 eValAlaHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 428
1246 ACGCGCGCGCGCGCGCGAGGAGCAAGAGC.....CT 1277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
429 ValSerAlaAspSerSerAlaSerMetAsnSerGlyValLeuValArg 445
1278 TCCCTGTTTGCGCGCGCTACGCGCTGCGCTGTGTGGGCTGTGAGG 1327
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
445 gProSerArgLeuSerSerSerGlyThrProMetLeuAlaGlyValSerG 462
1328 AGCATGGTCTCCGCGACGCGCCCGCAG 1353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
462 LysTyrrGlnLeuProGlnAspProArg 470

```

seq.name: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:US-08-451-822A-12

seq.documentation\_block:

```

: Sequence 12, Application US/08451822A
: Patent No. 5863888
:
: GENERAL INFORMATION:
: APPLICANT: Dionne, Craig A
: APPLICANT: Crumley, Greg
: APPLICANT: Jaye, Michael C
: APPLICANT: Schlessinger, Joseph
: TITLE OF INVENTION: Fibroblast Growth Factor Receptors
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESS: Rhone-Poulenc Rorer Legal Department
: STREET: 500 Arcola Road
: CITY: Collegeville
: STATE: PA

```

```

COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451.822A
FILING DATE: 26-MAY-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/323,430
FILING DATE: 14-OCT-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/934,372
FILING DATE: 21-AUG-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/549,587
FILING DATE: 06-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Sawitzky, Martin
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: A04966
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3816
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 822 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-451-822A-12

```

```

alignment_scores:
Quality: 423.50      Length: 442
Ratio: 1.580         Gaps: 14
Percent similarity: 60.633   Percent identity: 27.602

```

```

alignment block:
US-09-598-042A-2 x US-08-451-822A-12 ..

```

```

Align seg 1/1 to: US-08-451-822A-12 from: 1 to: 822

```

```

112 GGGCCGACATATGGGGCTGCATGCCACATGGAGGGAGCCGGCCGGCGG 161
    ||| :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
48 GlyAspLeuLeuGlnLeuArgCysArgLeuArgAspAspValGlnSer11 64
162 GACCATGATGGACCAAGATGCGCCGACCATTCACAGCGCTGGAGCCGCT 211
    :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
64 easn...TPrLeuArgAspLeuValGlnLeuAlaGlnSer...AsnArg79
212 TCCGCTGCTGCCGACAGGGCTGAAGGTGAAGCAGGTGAGCGGAGAGT 261
    |||:| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
79 hArGlnLeuThrGlnGlnGlnValGlnValGlnAspSerValProAlaAsp 95
262 GCGCGGCTGATGCTGTCGCAAGGCCACCAACGCGTTGGCAGCGTGAACG 311
    :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
96 SerGlnLeuValAlaCysValThrSerSerProSerGlnSer...AspTh 111
312 CAACCTACACCTCGCTGCTGATGATACATTAGCCGAGGAAGAGAGCC 361
    :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
111 rThrTyrrPheSerValAlaSerValSerAspAlaLeuProSerSerGlnAsp 128
362 TGGGCGCCACAGACTCTCTGGGGGTCAAGAGAGACCCCGCC..... 402
    |||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
128 sPrAspAspAspAspSerSerSerGlnGlnGlnValThrThrAspAsnThr 144
403 .....AGCCAGACATGGGACAGACCGCGCTTCACACAGCCCTCCCAAGAT 446
    :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|

```

```

145 LysProAsnArgMetProValAlaProTyrThrThrSerProGlnLysMe 161
447 GAGGCGCGGGTATGSCAAGCGCGCGGTGGTACCTGCGGCTGACAT 496
    |||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
161 tGlnLysLysLeuGlnAlaValProAlaAlaLysThrValLysPheLys 178
497 GCGTGGCGAAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 546
    || :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
178 ySerSerSerGlnThrProAsnProThrLeuArgThrLeuLysAsnGly 194
547 CAGCGCTTGACGCGCGCGAG.....GCGCGTGAAGCGCGCGCGCG 584
    :|| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
195 LysGlnPhe...LysProAsnArgMetLeuGlnGlnLysValAlaArg 210
585 GAAGAAGTGAACATGACCTGAAGAACTGCGGCGCGCGCGCGCGCGCG 634
    |||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
210 rAlaThrThrSerThrLeuMetAspSerValAlaProSerAspLysGly 227
635 AATACACCTGCGCGCTGTCGAACCGCGCGCGCGCGCGCGCGCGCGCG 684
    :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
227 snrTyrrCysLysLeuAlaGlnAsnGlnTyrLeuSerLeuAsnThrThr 243
685 AAGGTGATGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 734
    :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
244 GlnLeuAspValAlaGlnArgSerProAsnArgProLysLeuAlaGln 260
735 GCACCGCGTGAACACGAGCGTGGCTGGGGGACACACGCTCTCCAGT 784
    |||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
260 yLeuProAlaAsnLysThrValAlaLeuGlnGlnSerAsnValGlnPheMet 277
785 GCAAGTGGCGACGACGATGAAGCGCGCGCGCGCGCGCGCGCGCGCG 834
    |||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
277 yLysValTyrSerAspProGlnProGlnLysLeuThrLeuLysThr 293
835 GAGTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 884
    ||| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
294 GlnVal.....AsnGlnSerLysLysLeuLysProAsnLys 305
885 GAAGTTTGTG...GTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 931
    :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
305 uProTyrValGlnLeuLysThrAlaGlnVal.....AsnThrThr 320
932 GCGCTACCTGCAATGAAGCTGTGATGACCGCGCGCGCGCGCGCGCGCG 981
    :|| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
320 sPrLysGlnMetGlnValLeuGlnLysLeuArgAsnValSerPheGlnAla 336
982 GGCATGACATGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1031
    ||| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
337 GlnGlnTyrThrCysLeuAlaGlnAsnSerThrLeuLysThrAsnHis 353
1032 CGGCTTCTACCGCTGCTGCA.....GACCGAAACCGCGCGCGCGCG 1075
    |||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
353 rAlaThrPheThrValLeuGlnAlaLeuGlnGlnArgPro..... 366
1076 CTGTGGCTCTGCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1125
    :|| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
367 .....AlaValMetThrSerProLeuThrLeuGlnGlnLeu 378
1126 GCGATCCGACGCGCGCGCTGCTGATC.....CTGGGCGACCGCTG 1166
    :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
379 lIeTyrCysThrGlnAlaPheLeuLysSerCysMetValGlnSerVal 395
1167 CCTGTGGCTTTGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1197
    :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
395 eValTyrLysMetLysSerGlnThrLysLysSerAspPheHisSerGln 412
1198 ..TGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1245
    :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
412 eValAlaValHisLysLeuAlaLysSerThrProLeuArgGlnValThr 428
1246 ACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1277
    :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
429 ValSerAlaAspSerSerAlaSerMetAsnSerGlnValLeuLeuValAr 445

```



```

367 .....AlaValMetThrSerProLeuTyrLeuGluIleIle 378
1126 GGCATCCAGCCGGCGCTGCTTTCATC.....CTGGGACCCCTGCT 1166
379 lIeTyrCysThrGlyAlaPheLeuIleSerCysMetValGlySerVal11 395
1167 CCGTGGGCTTTCAGAGCCAGAGAGAGCCG..... 1197
395 eValIlyrIysMetIysSerGlyThrIlyIysSerIspHeIhSerGlm 412
1198 ..TGCACCCCGCGCGCTGCCCTGCCCTGGCGACCGCCCGCGCGGG 1245
412 eValIValhIstIysLeuAlaIlySerIleProLeuAArgArgIValThr 428
1246 AGGGCCCGGACCGCAGCGAGACAAGAC.....CT 1277
429 ValSerAlaAspSerSerAlaSerMetAsnSerGlyValLeuLeuValAr 445
1278 TCCCTGCTGGCGCGCTGACGCGCTGGCGCTGGGGGCTGTGTGAGG 1327
445 gProSerAlaGlySerSerSerGlyThrProMetLeuAlaGlyValSerG 462
1328 ACATGGGTCTCCGGCAGCCGCCAG 1353
462 lUtyrGluLeuProGluAspProArg 470

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-07-997-133-1
seq_documentation_block:
; Sequence 1, Application US/07997133
; Patent No. 528855
; GENERAL INFORMATION:
; APPLICANT: Bergonzoni, Laura
; APPLICANT: Mazue, Guy
; APPLICANT: Isacchi, Antonella
; APPLICANT: Roncucci, Romeo
; APPLICANT: Sarnientos, Paolo
; TITLE OF INVENTION: Extracellular Form of the Human
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/997,133
; FILING DATE: 28-DEC-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/642,755
; FILING DATE: 18-JAN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 528855man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 769-226-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OBAT UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 822 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

```

US-07-997-133-1

alignment_scores:
  Quality: 422.50      Length: 442
  Ratio: 1.582
  Percent Similarity: 60.407      Percent Identity: 27.602

alignment_block:
US-09-598-042a-2 x US-07-997-133-1 ..
Align seg 1/1 to: US-07-997-133-1 from: 1 to: 822

112 GCGCCGACTATGCGGCTGCGAGTGCAGTGAGAGGGAGCCCGCCGCT 161
111 ..... 111
48 GlyAspLeuLeuGlnLeuArgCysArgLeuAArgAspValGlnSer11 64
162 GACATGTGACCAAGACGCGCGGACCATCCAGACAGCGCTGAGCGCT 211
64 eAsn...TrrLeuAArgAspGlyValGlnLeu...AlaGluThrAsnArgT 79
212 TCCGCTGCTGCCGCGAGGGCTGAAGGTGAAGAGAGTGGAGCGAGAT 261
79 hrArgIleThrGlyGlnGluValGlnValGlnAspSerValProAlaAsp 95
262 GCGGCGGTACGTGTGCAAGCGCCACCAACGCGCTTGACGCTGAGCT 311
96 SerGlyLeuTyrAlaCysValThrSerSerProSerGlySer...AspTh 111
312 CAATTCACACCTCGTGCCTGATGATCATTTAGCCAGGAGGAGAGAGCC 361
111 rThrIlyrPheSerValAsnValSerAspAlaLeuProSerSerGluAspA 128
362 TGGGCGCCGACAGACTCTCTGCGAGGTCAAGAGAGCCCGCC..... 402
128 spAspAspAspAspAspSerSerSerSerGlnGlnLysGlnThrAspAsnThr 144
403 .....AGCCAGACAGTGGGACAGACCGCGCTTCACACACCCCTCCAGAT 446
145 LysProAsnArgMetProValAlaProTyrTrpThrSerProGluLysWe 161
447 GAGGCGCGGGGATGATGCGACGCGCGGTAGCTCCGCGGCTCAACT 496
161 tGluLysIlyLeuHnIstAlaValProAlaIlyStrValIlySpHeLysC 178
497 GCGTGGCCAGCGGGCACCTCGCGCCGACATCAGTGAAGAGAGAGC 546
178 ysProSerSerGlyThrProAsnProThrLeuAArgTrpLeuLysAsnGly 194
547 CAGGCGCTGACGCGCCAGAG.....GCCGCTGAGCCCGAGAA 584
195 LysGluPhe...LysProAspHisArgIleGlyGlyTyrValValArgTy 210
585 GAAGAAGTGACACTGAGCTGAAGACCTCGCGCGGAGAGAGAGAGGA 634
210 rAlaThrTrpSerIleIleMetAspSerValValProSerAspLysGlyA 227
635 AATPACCTGCGCGGTGTCAGACGCGGCGGCGCATCAACCCACCTAC 684
227 snTyrThrCysIleValGlnAsnGluTyrGlySerIleAsnHsthrTyr 243
685 AAGGTGATGTATCCAGCGGACCCGTTTCAAGCCGCTGCACAGGAC 734
244 GlnLeuAspValValGlnArgSerProHisArgProIleLeuGlnAla 260
735 GCACCCCGTGAACAGACAGCGTGACTGCGGGGAGACCACTGCTCCACT 784
260 yLeuProAlaAsnLysThrValAlaLeuGlySerAsnValGluPheMetC 277
785 GCAAGTGGCGAGCGACGACGTGAAGCCGCTGATGACGTGACGCGCGGT 834
277 yLysValIlyrSerAspProGlnProHisIleGlnTrpLeuLysHstIle 293

```

```

835 GAGTAGGCGCCGAGGCGCCGACAACTCCACATGATGTGGGCGCCA 884
      |||
294 GluVal.....AsnglySerLysAlaIleLysProAspSlnLe 305
885 GAAGTTGTG...GTGTCGCCGACGGGTGACGTGTGTCGCGCCGACG 931
      |||
305 uProTyrValGlnIleLeuLysThrAlaGlyVal.....AsnThrThra 320
932 GCTCCATCTCATTAAGCTCTGATACCCGTCGCGCCGACGAGATGCG 961
      |||
320 sPlysGluMetGluValIleuHisLeuArgAsnValSerPheGluAspAla 336
982 GGATGTACATCTGTCGTCGCGCCGACACATGCGGCTGACAGCTCCGAC 1031
      |||
337 GlyGluTyrThrCysLeuAlaGlyAsnSerIleGlyLeuSerHisSse 353
1032 CGCCTTCCTGACGCTGCTGCA.....GACCCAAACCGCCGAGGCGCAC 1075
      |||
353 rAlaThrPheThrValLeuGluAlaLeuGluArgPro..... 366
1076 CTGTGGCTCTGCTGCTGCGCCGACGTAGCTGCGCGCGCTGTGTCATC 1125
      |||
367 .....AlaValMetThrSerProLeuTyrLeuGluIle 378
1126 GCATCCGACGCGCGCTGTCTCATC.....CTGGGACACCTGCT 1166
      |||
379 IleTyrCysThrGlyAlaPheLeuIleSerCysMetValGlySerValII 395
1167 CCTGTGCTGTTCGACGCGCCGAGAAAGCCG..... 1197
      |||
395 eValTyrLysMetLysSerGlyThrLysSerAspPheHisSerLm 412
1198 ..TCGACCCCGCGCGCTGCGCCCTGCGCTGCGGACCGCGCGCGGG 1245
      |||
412 eValValHisLysLeuAlaLysSerIleProLeuArgArgGluValThr 428
1246 ACGCGCGCGCGACGCGGAGAGACAGAGAC.....CT 1277
      |||
429 ValSerAlaAspSerSerAlaSerMetAsnSerGlyValLeuValAr 445
1278 TCCCTGCTGGCGCGCGCTGACGCGCTGCTGTGGGCTGTGTAGG 1327
      |||
445 gProSerArgLeuSerSerGlyThrProMetLeuAlaGlyValSerG 462
1328 AGCATGGGCTCGCGGACGCGCGCGGAC 1353
462 IuTyrGluLeuProGluAspProArg 470

seq_name: /cgn2.6/plodata/2/1aa/5A_COMB.pep:us-08-166-717D-6
seq_documentation_block:
; Sequence 6, Application US/08166717D
; Patent No. 5789182
; GENERAL INFORMATION:
; APPLICANT: Yaron, Avner
; APPLICANT: Ornitz, David M.
; APPLICANT: Klagsbrun, Michael
; APPLICANT: Leder, Philip
; TITLE OF INVENTION: SYSTEM FOR ASSAYING BINDING
; TITLE OF INVENTION: TO A HEPARIN-BINDING GROWTH
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; OPERATING SYSTEM: WINDOWS 95

```

```

; SOFTWARE: WordPerfect (Version 7.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/166,717D
; FILING DATE: 12/14/93
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/631,717
; FILING DATE: 12/20/90
; ATTORNEY/AGENT INFORMATION:
; NAME: Kristina Bleker-Brady
; REGISTRATION NUMBER: 39,109
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 723-4123
; TELEFAX: (617) 723-6962
;
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 820
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-08-166-717D-6

```

```

alignment_scores:
  Quality: 413.50      Length: 441
  Ratio: 1.560         Gaps: 15
  Percent Similarity: 60.091  Percent Identity: 27.891

```

```

alignment_block:
US-09-598-042A-2 x US-08-166-717D-6 ..

```

```

Align seg 1/1 to: US-08-166-717D-6 from: 1 to: 820

```

```

112 GCGCGCATGTGCGGCTGCGAGTCCAGTGGAGGGGACCGCGCGCT 161
      |||
48 GlyAspLeuLeuGlnLeuArgCysArgLeuArgAspAspValGlnSerII 64
162 GACCATGTGGACCAAGATGAGC...GCGACCATTCACAGCGCTGAGGCC 208
      |||
64 eAsn...Trp*****AspIleValGlnLeuValGlnSer.....AsnA 78
209 GCTTCGCGGTGTCGCCGACGGGCTGAGGTGAAGCAGTGGAGCGGAG 238
      |||
78 rGThrArgIleThrGlyGluGluValGluValArgAspSerIleProAla 94
259 GATCCGCGCGCTGACGTCGTCGTCGATGACATTAGCCAGGGAAGAG 308
      |||
95 AspSerGlyLeuTyrAlaCysValThrSerTrpSerGlySer...As 110
309 CGTCACTACACCTCGTGTGTCGATGACATTAGCCAGGGAAGAG 358
      |||
110 pThrThrThrPheSerValAsnValSerAspAlaLeuProSerSerGluA 127
359 GCCTGGGCGCGACAGCTCTGTGGGGGTCAAG.....GACCC 399
      |||
127 sPaAspAspAspPheAspSerSerSerGluGluLysGluThrAspAsn 143
400 GCGAGCCAGCAGTGGGACGCGCGCTTCACACAGCGCTTCAGAGAG 449
      |||
144 ThrLysProAsnProValAlaProTyrThrPheSerProGluLysMetG 160
450 GCGCGGGGTATCGACGCGCGCTGAGTACCTCGTGGCTCAAGTGC 499
      |||
160 uLysLysLeuHisArgValProAlaAlaLysThrValLysPheLysCysP 177
500 TGGCAGCGGCGACCTGCGCGCGACATCACTGATGAAGAGGACGAC 549
      |||
177 roSerSerGlyThrProAsnProThrLeuArgThrLeuLysAsnGlyLys 193
550 GCCTTGACGCGCCAGAG.....GCCGCTGAGCCGACGAGAA 587
      |||

```



```

194 GluPhe...LysProAspHisArgIleGlyGlyTyrTyrValAlaGlyTyrAl 205
588 GAAGTGGACACTAGACTGAGAAGACTGGCGCGGAGACAGCGCAAT 637
||||| ||||| : : : : : ||||| : : : : :
209 atntrfserIleIleMetaspsrValAlaProSerAspArgSylant 226
||||| ||||| : : : : : ||||| : : : : :
638 ACACSTGCGCGGTGTGCAACCGCGCGGCGCGCATCAACCGCACTACAAG 687
||||| ||||| : : : : : ||||| : : : : :
226 ytrfCysIleValAlaGluAsnGlyTyrGlySerIleAsnHisThrTyrGln 242
||||| ||||| : : : : : ||||| : : : : :
688 GTGAGATGTGATCCAGGAGCGCCGTTCACAAACCCCGTGTCCACAGGCA 737
||||| ||||| : : : : : ||||| : : : : :
243 LeuAspValValGluArgSerProHisArgProIleLeuGlnAlaGlyLe 259
738 CCGCGTGAACACAGCGGTGGATCTCGGGGAGACCAAGCTCTTCCAGTCA 787
||||| ||||| : : : : : ||||| : : : : :
259 u***AlaAsnLysThrValAlaLeuGlySerAsnValAlaGluPheMetCysL 276
788 AGCTGGCGACCGACGTGAACCGCGGTATCCAGTGGCTGAACCGCGTGAAG 837
||||| ||||| ||||| ||||| : : : : : ||||| : : : : :
276 yValTyrSerAsp****ProHisIleGlnIleTyrLeuLysHisIleGlu 292
838 TAGCGGCGCGAGGGCGCGCAACACTGCACATGCATGTGGCGGCGCGCA 887
293 Val.....AsnGlySerLysIleGlyProAspAlaLeuP 304
888 GTTGTGTG...GTCTGCGCGACGGGTAGTGTGTGTGCGCGCGCGACGAGT 934
||||| ||||| ||||| ||||| : : : : : ||||| : : : : :
304 otyrValGlnIleLeuLysThrAlaGlyVal.....AsnThrThrSprL 319
935 CSTACSTCAATAAAGTGTCTATCAGCCGCGCGCGCGCGACAGATGCGGCG 984
: : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
319 ysrLysMetGluValLeuHisLeuAlaGlyAsnValSerPheGluAspAlaGly 335
985 ATGTACATATGCTGTGGCGCGCAACACCATGGCGTACAGCTTCCGACGCG 1034
||||| ||||| ||||| ||||| ||||| ||||| : : : : : ||||| : : : : :
336 GluTyrThrCysLeuAlaGlyAsnSerIleGlyLeuSerHisHisSerAl 352
1035 CTTCCTCACCGCTGTGCGCA.....GACCGAAACCGCGCGCGCGACCTGT 1078
||||| ||||| ||||| ||||| : : : : : ||||| : : : : :
352 atPheLeuThrValLeuGlnAlaLeuGlnGluAlaArgPro..... 364
1079 TGGCGCTCTGCTGCTGGCGCACTAGCTGCCGTGGCGCGCGGTCTATCGGC 1128
: : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
365 .....AlaValMetThrSerProLeuTyrLeuGlnIleLeu 377
1129 ATCCCAAGCGCGCGCTGTCTTCAAC.....CTGGCGACCGCTGCTCT 1169
||||| ||||| ||||| ||||| ||||| ||||| : : : : : ||||| : : : : :
378 TyrCysThrGlyAlaPheLeuIleSerCysMetLeuGlySerAlIleI 394
1170 GTGGCTTTTGGCAGGCGCGCAAGAACGCG.....T 1198
||||| ||||| ||||| ||||| : : : : : ||||| : : : : :
394 etyTyrLysMetLysSerGlyThrTyrLysSerAspPheHisSerIleMet 411
1199 GCACCGCGCGCGCTGCGCGCTGCGCTGCGCGCGCGCGCGCGCGGAGC 1248
||||| ||||| ||||| ||||| : : : : : ||||| : : : : :
411 ValAlaHisLysLeuAlaLysSerIleProLeuArgGlnValThrVal 427
1249 GCGCGGCGCGCGCAGCGGAGACAGACAGAC.....CTTCC 1280
||||| ||||| ||||| ||||| : : : : : ||||| : : : : :
428 SerAlaAspSerGluAlaSerMetLysSerGlyValAlaLeuLeuValArgPr 444
1281 CTGTGTTGGCGCGCGCTGACGCTGCGCGCGGTGTGTGGCGCGCTGTGAAGAGC 1330
||||| ||||| ||||| ||||| : : : : : ||||| : : : : :
444 oSerArgLysSerSerSerGlyThrProMetLeuAlaGlyValSerGluT 461
1331 ATGGGTCTCCGCGACGCGCGCGCGCAG 1353
: : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
461 yrcLysLeuProGluAspArg 468
seq_name: /cgn2_6/ptodata/2/laa/5A_COMB.pep:US-08-471-570-4
seq_documentation_block:

```

```

1 Sequence 4, Application US/08471570
2 Patent No. 5750371
3
4 GENERAL INFORMATION:
5 APPLICANT: IGARASHI, Koichi
6 APPLICANT: SENOO, Masaharu
7 APPLICANT: WATANABE, Tatsuya
8 TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
9 NUMBER OF SEQUENCES: 18
10
11 CORRESPONDENCE ADDRESS:
12 ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
13 ADDRESS: CUSHMAN
14 STREET: 130 Water Street
15 City: Boston
16 STATE: Massachusetts
17 COUNTRY: US
18 ZIP: 02109
19
20 COMPUTER READABLE FORM:
21 MEDIUM TYPE: Floppy disk
22 COMPUTER: IBM PC compatible
23 OPERATING SYSTEM: PC-DOS/MS-DOS
24 SOFTWARE: PatentIn Release #1.0, Version #1.25
25
26 CURRENT APPLICATION DATA:
27 APPLICATION NUMBER: US/08/471,570
28 FILING DATE: 06-JUN-1995
29 CLASSIFICATION: 435
30
31 PRIOR APPLICATION DATA:
32 APPLICATION NUMBER: US/08/149,664
33 FILING DATE:
34 APPLICATION NUMBER: US 07/743369
35 FILING DATE: 16-AUG-1991
36 ATTORNEY/AGENT INFORMATION:
37 NAME: LINK, Ernest V
38 REGISTRATION NUMBER: 29822
39 REFERENCE/DOCKET NUMBER: 40897
40 TELECOMMUNICATION INFORMATION:
41 TELEPHONE: (617)523-3400
42 TELEFAX: (617)523-6440
43 TELEX: 200291 STRE UR
44 INFORMATION FOR SEQ ID NO: 4:
45 SEQUENCE CHARACTERISTICS:
46 LENGTH: 526 amino acids
47 TYPE: amino acid
48 TOPOLOGY: linear
49 MOLECULE TYPE: protein
50 US-08-471-570-4

```

[illegible]

```

456 GGTGATTCGACGAGCCCGGTGGTATGCTCCGAGCGGCTCAATGGCTGGCCGA 508
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
457 gguhttslalaValProllaalaslntthValVslpheaUtyGsluPhe 67
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
506 GCGGGCACCCCTGGCCGCAATCACGTGATGAAGGACGACGACGCCCTTG 555
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67 lYgLIaSnProMetProHlnMetArgTlrPreuLysasnGlyGsluPhe 83
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
556 AGC.....CGCCACAGAGCCCGGTGACCCGACGAGAAGAAGTGGAC 596
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
84 LysGlnGlnHnStrGrleQlyGlyTrLysValArgAsnGlnHnStrPse 100
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
597 ACTGAGCCGTAAGAACCTGGCGGGGAGGACAGCGGCAATATACACCTGCC 646
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
100 rEuLlEmeGluSerValValProSerAspLysGlyAsnGlyThrCysV 117
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
647 GCGTGTGCAAGCGCGGGGCGCCATCAAGCCGCACTCAAGGTGGATGTG 696
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
117 alValGlnasnGlyTrGlySerLleAsnHnStrTrGlnHnLeAspVal 133
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
697 ATTCACGCGGACCCGTTCCAAAGCCCGTGTACAGAGCACGACCCCGTGA 746
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
134 ValGluArgSerProHnStrProIleLeuGlnAlaIolLeuProAlaAs 150
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
747 CAGAGCGGTGGACATTCGGGGGGGACACAGCTCTCCATGCGAAGTGGCGA 796
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
150 nAlaSerThrValValGlyGlyAspAlGluPheValCysLysValTrys 167
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
797 GCGACCTGAAGCCGCGTATCCACGTGCGTGAAGCGCGTGGAG..... 837
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
167 eRAspAlaGlnProHnStrIleGlnTrpLleLysHnValGlnLysasnGly 183
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
838 .....TACGGCGCCAGGCGCGCCACACTCCACACATCGATGTGGCGG 881
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
184 SerLysTrGlyGlyProAspGly..... 190
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
882 CCAGAGATTGTGTGCTGTGCCCAAGGCTGACGTG.....TGTCGCGGCG 925
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
191 .....LeuProTrGlyLeuLysValLeuLysHnStrIleTrGly 202
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
926 CGACGAGCTTCCTACCTCAATACAGCTGTGTACACCGCTGCCCGCCAGAC 975
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
202 leAsnSerSerAsnAlaGluValLeuAlaLeuPheAsnValThrGlnAla 218
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
976 GATCGCGGCGATGATCATCTTCCTGGCGCCACACATCGGCTACAGCTT 1023
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
219 AspAlaGlyGluTrGlyIleCysLysValSerAsnTrGlyIleGlyGlnAlaAs 235
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1026 CGCGACAGCGCTTCCTACACGCTGTGTGCCAGAACCCAAACCGCCAGGCGAC 1075
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
235 ngInSerAlaTrpLeuThrValLeuProLysGlnGlnAlaProGly... 250
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1076 CTGTGCGCTCTCTGTCTCGCGCACATGACCTGCGGTGGCCGTGATCATC 1125
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
251 .....ArgGlnLysGlnLleThrLAsaSerProAspTrGlyLeuGlnLle 264
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1126 GGCATCCCAACCGGCGCTGTCTTCATCTGGGACACCTGTCTGTGGCT 1175
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
265 AlaIleTrGlyCysIleGlyValAlaPheLleuIleAlaCysMetValThrVa 281
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1176 TTGCGCAGGCCCAAGAGAAGCCGTCGACCCCGCGGCT 1212
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
281 lIleLeuCysArgMetLysAsnThrTrHnLysLysPro 293
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

seq_name: /cgn2_6/ptodata/2/laa/5A_COMB.pep:US-08-471-570-10

seq_documentation_block:
: Sequence 10, Application US/08471570
: Patent No. 5750371
: GENERAL INFORMATION:
: APPLICANT: IGARASHI, Koichi
: APPLICANT: SENOO, Masaharu
: APPLICANT: WATANABE, Tatsuya

```

TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF  
 NUMBER OF SEQUENCES: 18  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &  
 ADDRESSEE: CUSHMAN  
 STREET: 130 Water Street  
 CITY: Boston  
 STATE: Massachusetts  
 COUNTRY: US  
 ZIP: 02109  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentln Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/471,570  
 FILING DATE: 06-JUN-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/149,664  
 FILING DATE:  
 APPLICATION NUMBER: US 07/743369  
 FILING DATE: 16-AUG-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: LINK, Ernest V  
 REGISTRATION NUMBER: 29822  
 REFERENCE/DOCKET NUMBER: 40897  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617)523-3400  
 TELEFAX: (617)523-6440  
 TELEX: 200291 STRE UR  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 652 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-471-570-10

```

alignment_scores:                                Length: 329
Quality: 379.50                                Gaps: 7
Ratio: 1.956
Percent Similarity: 58.967                    Percent Identity: 29.483

alignment_block:
US-09-598-042A-2 x US-08-471-570-10  ..

Align seg 1/1   to: US-08-471-570-10   from: 1   to: 652

265 GGCCTGTACGTTGCAAGGCCACCAAGCGCTTGCGACGCTGACGTC.. 312
|||||          :|||:  :|||:  :|||:  :|||:  :|||:  :|||:
5  GYARGPHELlCYLeuValValValThrMetAlaThrSerLeuVal 21
313 .....AACTACACCCCTGCTGCTGCTGATGACATTAGCCCAAGGAAG 355
|||||          :|||:  :|||:  :|||:  :|||:  :|||:  :|||:
21 aARGPProSerPheSerLeuVal..... 28
356 AGAGCGTTGGGGCCGACAGACGTCCTGCGGGGTCAAGAGACCCGCCAGC 405
|||||          :|||:  :|||:  :|||:  :|||:  :|||:  :|||:
29 .....GluAspThrThrLeu 33
406 CAGCAGTGGGACGACGCGCGCTTACACAGCCCTCCAGATGAGGCGCG 455
|||||          :|||:  :|||:  :|||:  :|||:  :|||:  :|||:
34 GlnProGlnGluAlaProTyrTrpThrAsnThrGlnIuysMetGlnIuysar 50
456 GGTGATGACGACGCGCCCGTGGGTAGTCCGCTCGGCTCAAGTCGGTCCA 505
|||||          :|||:  :|||:  :|||:  :|||:  :|||:  :|||:
50 GlnuHnAlaValAlProAlaAlaAsnThrValIuysPheArgIuysProAlaG 67
506 GCGGGACCGCTCGGCGCGACATCAGCTGATGAAGAGACGACAGCGCTTG 555
|||||          :|||:  :|||:  :|||:  :|||:  :|||:  :|||:

```

```

67  LYGlyAsnProMetProThiMetArgTrpLeuLysAsnGlyLysGluPhe 83
556  ACG.....CGCCAGAGGGCCCTGAGCCCGAGAGAGAGAGAGTGCAC 596
      |||      ::      |||:::      |||::
84  LysGlnGlnHisArgIleGlyGlyLysValArgAsnGlnHisTrpSe 100
597  ACTGAGCCCTGAGAAACCTCGCGCGGAGAGACAGCGGCAATACACCTGCC 646
      |||      :::      |||:::      |||:::      |||:::
100  rLeuLeuMetGluSerValValProSerAspLysGlyAsnTrpThrCysV 117
647  GCGTGTCCAGACCGCGCGGCGCCATCAACGCCCTACAGGTGATGTG 696
      |||:::      |||:::      |||:::      |||:::      |||:::
117  aValaGlnAsnGluTrpGlySerIleAsnHisThrTyrHisLeuAspVal 133
697  ATCCAGCGGACCCGTTCCAGGCCCTGCTCAGAGCGACGACCCCGTGA 746
      :::      |||:::      |||:::      |||:::      |||:::
134  ValGluArgSerProHisArgProIleLeuGlnAlaGlyLeuProAlaAs 150
747  CAGCAGGCGTCTCGGGGGGAGCCAGCCGCTCCAGTGCAGGTGCGCA 796
      |||:::      |||:::      |||:::      |||:::      |||:::
150  nAlaSerThrValValGlyGlyAspValGluPheValCysLysValTyrS 167
797  GCGAGCTGAAAGCCGGTGTATCCAGTGGCTGAAGCGGTGAG..... 837
      |||:::      |||:::      |||:::      |||:::      |||:::
167  eArgAlaGlnProHisIleGlnIleTrpIleLysHisValGluLysAsnGly 183
838  ....TACGGCGCGAGGGCGGCCCAACATCTCCACCATTGATGGCGG 881
      |||:::      |||:::      |||:::      |||:::      |||:::
184  SerLysTyrGlyProAspGly..... 190
882  CCAGAGATTGTGTGCTGCTGCCCGACGGGTGACGTG.....TGGTGGCGG 925
      |||:::      |||:::      |||:::      |||:::      |||:::
191  ....LeuProTyrLeuLysValLeuLysHisSerGlyI 202
926  CCGAGCGCTCTACCTCATATAAGTGTCTATCACCCTGCCCGCAGAG 975
      |||:::      |||:::      |||:::      |||:::      |||:::
202  LeAsnSerSerAsnAlaGluValLeuAlaLeuPheAsnValThrGluAla 218
976  GATGGGGGATGTACATCTGCTTGGCGCCCAACACCATTGGCTTACAGCTT 1025
      |||:::      |||:::      |||:::      |||:::      |||:::
219  AspAlaGlyGluTyrIleCysLysValSerAsnTyrIleGlyGlnAlaAs 235
1026  CCGCAGCGGCTCTCTACCGTGTGCCAGACCCAAACCGCGCGGCCAC 1075
      |||:::      |||:::      |||:::      |||:::      |||:::
235  nGlnSerAlaTrpLeuThrValLeuProLysGlnGlnAlaProGly.... 250
1076  CTGTGGCTCTCTGTCTCTGCGCCATAGCCCTGCCGCGCTGGTGCATC 1125
      |||:::      |||:::      |||:::      |||:::      |||:::
251  ....ArgGluLysGluIleThrAlaSerProAspTyrLeuGluIle 264
1126  GGCATCCAGCGCGCGCTGTCTTTCATCTGGGACACCCTGCTGTGCT 1175
      |||:::      |||:::      |||:::      |||:::      |||:::
265  AlaIleTyrCysIleGlyValPheLeuIleAlaCysMetValValThrVa 281
1176  TTGCCAGCGCCAGAGAACCTGCACCCCGCGCT 1212
      |||:::      |||:::      |||:::      |||:::      |||:::
281  IleLeuCysArgMetLysAsnThrLysLysPro 293
seq_name: /cgn2_6/ptodata/2/lae/5B_COMB pep:US-08-977-767-3
seq_documentation_block:
: Sequence 3, Application US/08977767
: Patent No. 5972684
: GENERAL INFORMATION:
: APPLICANT: Bandman, Olga
: APPLICANT: Yue, Henry
: APPLICANT: Greenwald, Sara
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: CARBONIC ANHYDRASE VIII
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive

```

```

: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/977,767
: FILING DATE: Herewith
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PP-0423 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-855-0555
: TELEFAX: 650-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1345 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: Genbank
: CLONE: 1532042
: US-08-977-767-3

```

```

alignment_scores:
  Quality: 371.50      Length: 582
  Ratio: 1.446        Gaps: 31
  Percent Similarity: 44.158      Percent Identity: 25.601

```

alignment\_block:

US-09-598-042a-2/rev x US-08-977-767-3 ..

Align seg 1/1 to: US-08-977-767-3 from: 1 to: 1345

```

2639  ACACGACTGACAGGCTTTCACGGCACACCTCCCAAAAGACGTCATCTG 2590
      || |||:::      |||:::      |||:::      |||:::      |||:::
801  ThrGlyThrGlyGlyCysThrGly.....GlyThrAlaGly 813
2589  CGTGTGTCCAGCAGCATATCTGCACCTGTGTGTCATGTGTGTCGGAA 2540
      |||:::      |||:::      |||:::      |||:::      |||:::
813  ThrCysAlaAlaAlaCysCysGlyThrThrCysCysThrCysAlaG 830
2539  GTGTGTGCCAGCAGCATATCTGCATGTGTGTCGTCGTGTATCCGA 2490
      |||:::      |||:::      |||:::      |||:::      |||:::
830  LysAlaCys...GlyCysCysThrCysCysThrThrAlaCysCysGly 845
2489  CAGCATATCTGCACGTCGTGCATGTCCAGACAGCATATCTGTGCATG 2440
      |||:::      |||:::      |||:::      |||:::      |||:::
846  ThrGlyAlaCysAlaCys.....CysAlaThrCysAlaCys 857
2439  TGTGTGTCCAGCAGCATATCTGCACGTCGTGTGAGTGTGAGCAGCA 2394
      |||      |||:::      |||:::      |||:::      |||:::
857  scCysCysGlyCysAlaThrCysThrCysCysThrAlaThrAlaAlaGly 874
2393  ....TTATCTGTGTGTGTGTCCAGCAGCATATCTGCGTCG... 2357
      ::      |||      |||      |||      |||      |||
874  AlaAlaThrGlyAlaThrGlyCys.CysThrAlaCysThrThrCysTh 890
2356  ....GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2330
      |||:::      |||:::      |||:::      |||:::      |||:::
890  ThrCysAlaAlaGlyAlaCysCysThrThrAlaGlyCysCysThrGlyG 907

```

```

2232 .....GCATGGCTGTGTGTGCGGACAGCATGTCTGGTGTGT ..... 2291
      :::::::::::::: ::::: ||| -|||
907  LyaIaGlyCysThrCysCysThrIyThrIyCysCysCysGlyAla 923
2290 ...GTGTAGTACAGCATATCCGCGTGTGTCCAGGACAGCATATCCG 2244
      :::::::::::::: ::::: ||| -|||
924  GlyThrcysCysThrThrIyGlyCysThrThrcysAla..ThrcysAla 940
2243 CCTGTGACGCTGTCTGGAAGTGTGTGTGTGTGTGTGTGTGTGTGT 2194
      :::::::::::::: ::::: ||| -|||
940  IacysCysThr.....AlaThrcysAlaGlyGlyCysThrCys 953
2193 GTGTGCTGTGTCCAGACAGCATATCTGTGTGTGTGTGTGTGTGTGTGT 2144
      :::::::::::::: ::::: ||| -|||
954  ThrcysThrcysAlaGlyCysAlaCysCys..CysCysAlaCysCysCys 970
2143 CAGGCAATATCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2094
      :::::::::::::: ::::: ||| -|||
970  hrgly.....CysThrcysGlyGlyAlaGlyAlaCys..... 980
2093 AGGACATATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2044
      :::::::::::::: ::::: ||| -|||
981 .....ThrcyThrThrcyAlaCys 986
2043 GTGGCTCCAGACAGCATATCCGTGTGTGTGTGTGTGTGTGTGTGTGT 1994
      :::::::::::::: ::::: ||| -|||
986  scysThrIyGlyAlaThrcysCysThrCysAlaThrcyThrcyAlaCysAla 1003
1993 CTGGCATATGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1944
      :::::::::::::: ::::: ||| -|||
1003  IagIyGlyGlyCysCysCysThrcysAlaAlaThrcyAlaCys 1019
1943 CTGGCATGTCTGTGT ..... 1929
      :::::::::::::: ::::: ||| -|||
1020  CysThrcysCysThrcysCysAlaGlyCysThrcyAlaThrcyAlaThrcy 1036
1928 .....GGTGTGCGGTGCATTTGGGCGGCGGCGGCGGCGGCGGCGG 1901
      :::::::::::::: ::::: ||| -|||
1036  rcysCysCysThrcyThrcyCysAlaCysThrThrcyThrcyAlaGlyCys 1053
1900 TGTGTGCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1869
      :::::::::::::: ::::: ||| -|||
1053  yscysAlaGlyAlaAlaThrcysCysThrcysCysGlyGlyThrcys 1069
1868 GTGTGTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1835
      :::::::::::::: ::::: ||| -|||
1070  CysCysAlaGlyThrcysThrThrcyCysAlaGlyAlaGlyCysThrCys 1086
1834 TGCGCGCATGTGT ..... 1822
      :::::::::::::: ::::: ||| -|||
1086  salagIyCysCysGlyThrcyAlaAlaCysGlyCysGlyCysGlyCysCys 1103
1821 .....GTGCATATATCATCCAGTGTGTGTGTGTGTGTGTGTGTGT 1786
      :::::::::::::: ::::: ||| -|||
1103  yscysCysThrIyCysAlaGlyCysCysCysThrThrcyCysGlyGly 1119
1785 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1736
      :::::::::::::: ::::: ||| -|||
1120  CysAlaCysAlaGlyGlyGlyCysCysThrThrcyAlaGlyGlyGlyGly 1136
1735 CCATTTCTCTCCGCGCATATGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1686
      :::::::::::::: ::::: ||| -|||
1136  yCysAlaAlaCysAlaGlyAlaCysCysCys.....C 1149
1685 CATCTCCCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1636
      :::::::::::::: ::::: ||| -|||
1149  ysgIyGlyCysCysAlaCysCysGly.....AlaGly 1160
1635 ACAGTGTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1586
      :::::::::::::: ::::: ||| -|||
1161  AlaGlyGlyGlyGlyThrcyGlyGlyAlaGlyGlyGlyCys..... 1175

```

```

1585  ACGTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1536
      ||| :::::::::::::: ::::: ||| -|||
1176 .....CysCysAlaAlaCysThrAlaCysCysGlyCys..Cys 1187
1535 GTAGAGTTGGGGTACACTTAGGGCCAGCACTGG...CTGGGCGCA 1489
      :::::::::::::: ::::: ||| -|||
1188  Thr..GlyCysAlaThrcyThrcyGlyAlaThrcyAlaThrcyThrcy 1204
1488 GTAGTGTGTGGGGGCGGCGGAGACCCATCTGTGTGTGTGTGTGTGT 1439
      ||| :::::::::::::: ::::: ||| -|||
1204  yscysCysCysAlaThrcyGlyThrcyThrcyCys..ThrcyAlaGly 1220
1438 CAATACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1389
      :::::::::::::: ::::: ||| -|||
1220  yAlaCysThrcysCysCysAlaThrcyGlyAlaGlyAlaGlyAla..Gly 1236
1388 .....TTAGGCGCAGCACTGGGCGGCGGCGGCGGCGGCGGCGGCG 1345
      ||| :::::::::::::: ::::: ||| -|||
1237  ThrcysCysGlyCysCysThrcyGlyCysCysThrcyThrcyCysAla 1253
1344 TGCGGAGACCCATGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1295
      ||| :::::::::::::: ::::: ||| -|||
1253  aAlaGly.....CysCysThr..CysCys..... 1260
1294 GGGCGGCCAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1245
      ||| :::::::::::::: ::::: ||| -|||
1261 .....CysCysAlaCysCysGlyAlaGlyGlyGly 1270
1244 CCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1214
      :::::::::::::: ::::: ||| -|||
1270  yGlyAlaGlyGlyGlyGlyAlaGlyThrcyGlyCysAlaCysCysCysCys 1287
1213 .....CAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1181
      ||| :::::::::::::: ::::: ||| -|||
1287  ysaAlaAlaAlaAlaCysAlaAlaAlaGlyCysThrAlaThrcy..... 1301
1180 GGCAAGCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1151
      :::::::::::::: ::::: ||| -|||
1302  AlaAlaAlaGlyGlyGlyAlaCysAlaGly 1311
seq_name: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:us-07-921-807B-5
seq_documentation_block:
: Sequence 5, Application US/07921807B
: Patent No. 5474914
: GENERAL INFORMATION:
: APPLICANT: SPAETE, RICHARD
: TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
: TITLE OF INVENTION: OF VIRAL PROTEINS
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CHIRON CORPORATION
: STREET: 4560 Horton Street - R440
: CITY: Emeryville
: STATE: CA
: COUNTRY: USA
: ZIP: 94608-2916
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/921,807B
: FILING DATE: 29-SEP-1992
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: MCCLUNG, BARBARA G.
: REGISTRATION NUMBER: 33,113
: REFERENCE/DOCKET NUMBER: 0209.001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (510) 601-2708
: TELEFAX: (510) 655-3542

```











## alignment\_scores:

Quality: 363.00 Length: 367  
 Ratio: 1.643 Gaps: 11  
 Percent Similarity: 60.218 Percent Identity: 28.338

## alignment\_block:

US-09-598-042a-2 x US-07-921-807B-6 ..

Align seg 1/1 to: US-07-921-807B-6 from: 1 to: 733

```

337 GACATTAGCCAGGAGAGAGACCTGGGCGCCAGACAGCTCTCTGGGG 386
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
31 AspalaleuProserSergluAspaAspaAspaAspaSerserSerge 47
387 TCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
47 rglugluysglutThrAspaThrLysProAsnArgMetProValAlaP 64
422 CGCGCTTACACAGCCCTCCAGATGAGGCGCGGATCGACAGCGGCC 471
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
64 rolyrThrPntSerProgluLysMetgluLysLysLeuHisAlaValPro 80
472 GTGGTAGCTCCGCGGCTCAAGTCGTCGGCCAGCGGCGCCCTGGCC 521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
81 AlaAlaLysThrValLysPheLysCysProSerSerglyThrProAsn 97
522 CGACATCACGTGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 567
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
97 oThrLeuArgTyrPheLysAsnLysgluPhe...LysProAspHisA 113
568 .....GCCGCTGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 609
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
113 rglleuglytyrLysValArgTyrAlaThrThrPntSerlelleMetAsp 129
610 AACCTCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 659
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
130 SerValValProSerAspLysgluAsnTyrThrCyslleAlaLysnGl 146
660 CGCGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 709
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
146 uTyrgLysSerleAsnHisThrTyrGlnLeuAspValAlaGluArgSer 163
710 GTTCCAGAGCCGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 759
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
163 roHisArgProIleuGlnAlaGlyLeuProAlaAsnLysThrValAla 179
760 TTGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 809
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
180 LeuGlySerAsnValGluPheMetCysLysValTyrSerAspProGlnP 196
810 GGTGATTCAGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 859
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196 oHislleGlnThrPheLysHislleGluVal.....AsnG 208
860 ACTCCACCATCATGTGGGCGGCGGAGAGAGAGAGAGAGAGAGAGAG 906
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
208 LysSerLyslleGlyProAspAsnLeuProTyrValGlnlleLysThr 224
907 GGTGAGCTGTGGTGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 956
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
225 AlaGlyVal.....AsnThrThrAspLysgluMetGluValLeuHis 239
957 CACCGCTGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1006
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
239 uArgAsnValSerPheGluAspAlaGlyLysTyrThrCysLeuAlaGlyA 256
1007 ACACCATGGGCTACAGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1053
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
256 snSerlleGlyLeuSerHisHisSerAlaThrPheThrValLeuGluAla 272
1054 ....GACCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

273 LeuGluGluArgPro.....AlaValMetTh 281
1101 TAGCCGTGCGCGCGGCTGCTCATCGGCATCCAGCGGCGGCTGTCTCA 1150
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
281 rSerProLeuTyrLeuGluIlellelleTyrCysThrGlyAlaPheLeu 298
1151 TC.....CTGGACACCTTCCTCTGCTGGCTTTCAGAGAGAGAGAG 1191
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
298 LeSerCysMetValGlySerValIleValTyrLysMetLysSerglyThr 314
1192 AACCG.....TGCACCCCGCGGCTGCGCTCC 1220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
315 LysLysSerAspPheHisSerGlnMetAlaValHisLysLeuAlaLys 331
1221 CCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1270
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
331 rleProLeuArgArgGluValThrValSerAlaAspSerSerAlaSer 348
1271 AGGAC.....CTTCCCTGCTTGGCGCGGCTGCGCTCAGAGAGCT 1302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
348 eLAsnSerglyValLeuLeuValArgProSerArgLysSerSergly 364
1303 GGCCTGTGTGTGGGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
365 ThrProMetLeuAlaGlyValSerGluTyrGluLeuProGluAspPro 381
1353 G 1353
381 g 381
seq_name: /cgn2_6/ptodata/2/iae/5A_COMB.pep:us-08-441-944a-6
seq_documentation_block:
; Sequence 6, Application US/08441944A
; Patent No. 5767250
; GENERAL INFORMATION:
; APPLICANT: SPAETE, RICHARD
; TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
; TITLE OF INVENTION: OF VIRAL PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: 4560 Horton Street - R440
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/441,944A
; FILING DATE: 16-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/921,807
; FILING DATE: 29-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCUNG, BARBARA G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0209,001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 733 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

256 snSerIleGlyLeuSerHisHisSerAlaTrpLeuThrValIleuGluAla 272

; TOPOLOGY: 1.1



TELEX: 200291 STRE UR  
 INFORMATION FOR SEQ ID NO: 12:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 240 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-471-570-12

## Alignment\_scores:

Quality: 362.50 Length: 237  
 Ratio: 2.401 Gaps: 4  
 Percent Similarity: 63.713 Percent Identity: 34.599

## alignment\_block:

US-09-598-042A-2 x US-08-471-570-12 ..

Align seg 1/1 to: US-08-471-570-12 from: 1 to: 240

```

391 GAGGACCCGCCAGCCAGCTGGGACAGCCGCGCTTCACACAGCCCTC 440
||||| : : : : : ||| : : : : :
9 GluSprThrThrLeuGluProGluGlyAlaProTyrTrpThrAsnThrG1 25
441 CAAGATGAGCGCGCGGTGATCGACAGCGCGCGGTAGCTCCGCGGC 490
||||| : : : : : ||| : : : : :
25 uLysMetGluLysArgLeuHisAlaValProAlaAlaAsnThrValLysP 42
491 TCAAGTGGCGTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
||||| : : : : : ||| : : : : :
42 heaGCGSprProAlaGlyGlyAsnProMetProThrMetArgTrpLeuLys 58
541 GAGCAGCAGCGCGCTTGAGC.....CGCCAGAGAGCGCGCGCGCGCG 581
|| : : : : : || : : : : :
59 AsnGlyLysGluProPheLysGlnGlnHisArgIleGlyTyrLysValAr 75
582 GAAGAAGAGTGGACACTGAGCCCTGAAGAAGCTCGCGCGGAGACAGCG 631
||||| : : : : : ||| : : : : :
75 gAsnGlnHisTrpSerLeuIleMetGluSerValValProSerAspLysG 92
632 GCAATATACACTGCGCGGTGTCGAACCGCGCGCGCGCGCGCGCGCGCG 681
||||| : : : : : ||| : : : : :
92 LysAsnTyrThrCysValValGluLysnGluTyrGlySerIleAsnHisThr 108
682 TACAAGGTGATGATTCACAGCGCGCGCTTCCAGCGCGCGTGTACAGCG 731
|| : : : : : || : : : : :
109 TyrHisLeuAspValValGluArgSerProHisArgProIleLeuGlnAl 125
732 CAGCCACCGCGCGTGAACACAGCGGTGGACTTCGGGGGACACACGCTTCC 781
||||| : : : : : ||| : : : : :
125 aGlyLeuProAlaAsnAlaSerThrValValGlyLysAspValGluPheV 142
782 AGTCAAGGTGCGCAGCGCGCGTGAAGCGCGGTGATCCAGTGGCTGAAGCG 831
||||| : : : : : ||| : : : : :
142 aLysLysValTyrSerAspAlaGlnProHisIleGlnTrpIleLysHis 158
832 GTGGAG.....TACGGCGCGGAGGCGCGCGCACAACTCCAC 866
||||| : : : : : ||| : : : : :
159 ValGluLysAsnGlySerLysTyrGlyProAspLys..... 170
867 CATGAGTGTGGCGCGCGCGAGATTGTGTGTCGCGCGCGCGGTGACTG 915
||||| : : : : : ||| : : : : :
171 .....LeuProTyrLeuLysVal 177
916 .....TGGTGGCGCGCGCGCGCTACTCAATAGCTGTCATACCG 960
||||| : : : : : ||| : : : : :
177 euLysHisSerGlyIleAsnSerSerAsnAlaGluValLeuAlaLeuPhe 193
961 CGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1010
||||| : : : : : ||| : : : : :
194 AsnValThrGluAlaAspAlaGlyGluTyrIleCysLysValSerAsnTy 210
1011 CATTGGGCTAAGCTTCGCGAGCGCGCTTCCTACCGCTGTCCAGACCCAA 1060

```

```

210 rIleGlyGlnAlaAsnGlnSerAlaTrpLeuThrValLeuProLysGlnG 227
1061 AACCGCCAGCG 1071
227 LnaAlaProGly 230

```

seq\_name: /cgn2\_6/prodata/2/laa/5A\_COMB.pep:US-08-070-165F-6

## seq\_documentation\_block:

```

Sequence 6, Application US/08070165F
Patent No. 5750365
GENERAL INFORMATION:
APPLICANT: Chiu, Ing-Ming
TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESS: Ing-Ming Chiu
STREET: S2052 Davis Medical Research Center, 480 West
STREET: 9th Avenue
CITY: Columbus
STATE: Ohio
COUNTRY: USA
ZIP: 43210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/070,165F
FILING DATE:
CLASSIFICATION: 435
TELECOMMUNICATION INFORMATION:
TELEPHONE: (614)-293-8093
TELEFAX: (614)-293-5631
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 729 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-070-165F-6

```

## alignment\_scores:

Quality: 361.50 Length: 316  
 Ratio: 1.817 Gaps: 11  
 Percent Similarity: 62.975 Percent Identity: 30.063

## alignment\_block:

US-09-598-042A-2 x US-08-070-165F-6 ..

Align seg 1/1 to: US-08-070-165F-6 from: 1 to: 729

```

268 GTGTAGCTGTGCAAGCGCGCGCTTGGCAGCGCGCTGACCTCACTA 317
||| : : : : : ||| : : : : :
11 ValMetValAlaThrAlaThrLeuSerLeuAlaArgProSerTyrAsn.. 26
318 CACCTTCGTCGCTGCGTGAATGACATTAGCCAGGAAGAGAGCGTGGCG 367
||| : : : : : ||| : : : : :
27 .....IleAlaGluAsp.....ThrThrLeuGluP 35
368 CGGAC.....AGCTCTCGGGGCGTCAAGAGGAC.....CCGGCGAC 405
||||| : : : : : ||| : : : : :
35 roGluAspAlaAsnSerSerGlyAspAspGluAspAspAsnAspGlySer 51
406 CAGCAGTGGGCGCAGA.....CCGCGCTTCACACA 434
||||| : : : : : ||| : : : : :
52 GluAspPheThrAsnAspAsnHisMetArgAlaProTyrTrpThrAs 68
435 GCCCTCAAGATGAGCGCGCGGTGATCGCAGCGCGCGGTGGTGGTCCG 484

```



```

152 H1STHrYrH1sLeuAspValValGluArgSerProH1sArgProH1leLeu 168
726 CACAGGCGACGACCCCGTGAACACGAGCGGTGACTTCGGGGGACACCT 775
168 uGlnAlaGlyLeuProAlaAsnThrThrThrValGlyGlyAspAlaG 185
776 CCTTCAGTCAAGGTGGCGACGACGAGCTGAGCCGGTGTATCAGTGGG 825
185 LuPheValCysLysValTyrSerAspAlaGlnProH1sIleGlnTrpIle 201
826 AAGCCGCTGAGTACGGCGCGCGGCGCCACAACTCCACCATGATGT 875
202 ArgHisPheGluLeu.....AsnGlySerLysIleGlyPyr 213
876 GGGCGCGCGAGATTGTG...GTGCTGCCCGCGGTGAGCGGTGGTGGC 922
213 CAspGlyHisProTyrLeuLysValLeuLysAlaAlaGlyAl.....A 228
923 GGGCGCGAGGCTCTACCTCAATAAGCTGCTCATCACCCTGCGCGCCAG 972
228 snThrThrAspLysGluIleGluValLeuTyrValArgAsnValSerPhe 244
973 GACGATGGGGGATATACATCGCTTGGCGCCACACCATGCGCTACAG 1022
245 GluAspAlaGlyGluTyrThrCysLeuAlaGlyAsnSerThrGlyIleSe 261
1023 CTTCGCGACGCGCTCTCACCGCTGCTGCTGCTGCGCCAGCCCAACCGCAGGCG 1072
261 rTyThrIsthrAlatrpLeuThrValLeuProAspGlu.....A 275
1073 CACCTGTGGCTCTCTGTCCTGGCGACTAGCTGCGCGCTGGCGCTGTC 1122
275 rGluLeuAspSerSerSerGluTyrThrGluIle..... 286
1123 ATGGGATCCAGCGGCGGTGTCTCTGCTGCTGCGCGACCTGCTCTG 1170
287 .....AlaIleTyrCysValAlaGlyIlePheLeuIle 296
seq_name: /cgn2_6/prodata/2/1aa/3A_COMB.pep:us-07-640-029-3
seq_documentation_block:
; Sequence 3, Application US/07640029
; Patient No. 5228501
; GENERAL INFORMATION:
; APPLICANT: Kiefer, Michael C.
; APPLICANT: Valenzuela, Pablo D.T.
; APPLICANT: Barr, Philip J.
; TITLE OF INVENTION: Expression and Use of Human Fibroblast
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/640,029
; FILING DATE: 19910111
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: CH-165
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2708
; TELEFAX: 510-655-3542

```

```

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 729 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-640-029-3

alignment_scores:
  Quality: 360.00      Length: 362
  Ratio: 1.651        Gaps: 10
  Percent Similarity: 60.221  Percent Identity: 29.006

alignment_block:
US-09-598-042A-2 x US-07-640-029-3

Align seg 1/1 to: US-07-640-029-3 from: 1 to: 729

337 GACATTAGCCAGGAGAGAGAGCTGGGCGCGACAGCTCTCTGGAGG 386
111
31 AspAlaLeuProSerSerGluAspAspAspAspAspAspSerSerSe 47
387 TCAAGAG.....GACCCCGCGACGACGAGTGGGCGACGCGGCGCT 427
47 rGluGluLysGluThrAspAsnThrLysProAsnProValAlaProTyrT 64
428 TCACAGAGCCCTCCAGATGAGGCGCGGCTGATCGACGCGCCGCTGGGT 477
111
64 rPThrSerProGluLysMetGluLysLysLeuHisAlaValProAlaAla 80
478 AGCTCCGTCGCGCTCAAGTGCCTGCGCGCGGCGGCGACCTGCGCGCAT 527
111
81 LysThrValLysPheLysCysProSerSerGlyThrProAsnProThrIle 97
528 CACGTGATG...AAGAGCAGACGAGGCTTGACGCGCGCCAGAGCGCTG 574
111
97 uArgTrpIleGlyLysGluPheLysProAspHisArgIleGlyGlyTyrL 114
575 AGCCGAGGAGAGAGAGAGTGAAGTGAAGCTGAGAGCGGAG 624
111
114 yValArgTyrAlaThrTrpSerIleIleMetAspSerValAlaProSer 130
625 GACAGCGCAATACACCTGCGGCTGCGAGCGCGCGCGCGCCCAACAA 674
111
131 AspLysLysAsnTyrThrCysIleValGluAsnGluTyrIleSerIleAs 147
675 CGCCACCTACAGGTGATGTGATCCAGCGGACCGCTTCCAGCGCGTGC 724
111
147 H1stThrTyrGlnLeuAspValValGluArgSerProH1sArgProH1le 164
725 TCACAGGCGACGACCCCGTGAACACGAGCTGAGCTTCGGGGGACGACG 774
111
164 euGlnAlaGlyLeuProAlaAsnLysThrValAlaLeuGlySerAsnVal 180
775 TCCTTCAGTCAAGGTGCGCGAGCGAGTGAAGCCGGTATCCAGTGGCT 824
111
181 GluPheMetCysLysValTyrSerAspProGlnProH1sIleGlnTrpIle 197
825 GAAGCGGTGAGTACGGCGCGCGGCGCCACAACTCCACCATGATG 874
111
197 uLysHisIleGluVal.....AsnGlySerLysIleGlyPyr 209
875 TGGGCGCGCGAGATTGTG...GTGCTGCCCGCGGTGACGCTGTGCTG 921
111
209 rAspAsnLeuProTyrValGlnIleLeuLysThrAlaGlyAl..... 223
922 CGGCGCGAGGCTCTACCTCAATAAGCTGCTCATACCGCGTCCCGCCA 971
111
224 AsnThrThrAspLysGluMetGluValLeuHisLeuAlaArgAsnValSerP 240
972 GAGCATGGCGGCGATGTACATCTGCTTGGCGCGCAACACCATGCGCTACA 1021

```

```

240 egluaspalaglyglutylthrcysleualaglyasnserileglyleus 257
      ::::::::::: ||| ::::::::::: |||
1022 GCTTCGGCAGCGCTTCTCTACCGTGTCCCA.....GACCCAAAACCG 1065
      || ::::::::::: ||| ::::::::::: |||
257 exHisHisserAlaIerPleuThrValleuGluAlaIeuGluIuArgPro 273
1066 CCAGGGCCACTGTGGCTTCCTCGCTCGGCGACCTAGCTGCGCGGCC 1115
      ::::::::::: ||| ::::::::::: |||
274 .....AlaValMetThrSerProleuTyrlle 282
1116 CGNGTCATCGCATCCAGCGGCTGTCTTCATC.....CTGG 1156
      ::::::::::: ||| ::::::::::: |||
282 ucluielleIleTyrcysThrGlyAlaIleuIleSerCysMetValG 299
1157 GCACCTGTCTCTGTGTGGCTTGTGGCAGCCAGAGAGCG..... 1197
      ::::::::::: ||| ::::::::::: |||
299 lYserValIleValIyrlYsMetlySerGlyThrlYslySerAspPhe 315
1198 .....TGCACCCCGCGCTGCGCTCCCTCCCTGCTGGGACCG 1235
      ::::::::::: ||| ::::::::::: |||
316 HisSerGluMetAlaValHislySleuAlaIySerIleProleuArgAr 332
1236 CCGCGCGGGAGCGCGCGCGCGCGAGAGAGAGAGAC..... 1275
      ::::::::::: ||| ::::::::::: |||
332 gclnValThrValSerAlaAspSerSerAlaSerMetAsnSerGlyValL 349
1276 .....CTTCCCTGTGTGGCGCGCGCTGCGCGCTGCTGGG 1317
      ::::::::::: ||| ::::::::::: |||
349 euLeuValArgProSerArgIleuSerSerSerGlyThrProMetLeuAla 365
1318 CTGTGTGAGAGCAGCGGTCTCCGCGACGCCCGCCAG 1353
      ::::::::::: ||| ::::::::::: |||
366 GlyAlaSerGluTyrlGluLeuProGluAspProArg 377
seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-08-070-165F-10
seq_documentation_block:
: Sequence 10, Application US/08070165F
: Patent No. 5750365
: GENERAL INFORMATION:
: APPLICANT: Chiu, Ing-Ming
: APPLICANT: Poulin, Matthew L
: TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Ing-Ming Chiu
: STREET: S2052 Davis Medical Research Center, 480 West
: STREET: 9th Avenue
: CITY: Columbus
: STATE: Ohio
: COUNTRY: USA
: ZIP: 43210
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/070.165F
: FILING DATE:
: CLASSIFICATION: 435
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (614)-293-8093
: TELEFAX: (614)-293-5631
: INFORMATION FOR SEQ ID NO. 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 731 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-070-165F-10

```

```

alignment_scores:
      Quality: 339.50      Length: 317
      Ratio: 1.741      Gaps: 11
      Percent Similarity: 61.514      Percent Identity: 29.338
alignment_block:
US-09-598-042a-2 x US-08-070-165F-10 ..
Align seg 1/1 to: US-08-070-165F-10 from: 1 to: 731
268 GTTAGCTAGCTCAAGGACGACCAAGCGGCTTGCAGCGCTGAGCTCACTA 317
      ||| ::::::::::: ||| ::::::::::: |||
11 ValMetValAlaThrAlaThrLeuSerLeuAlaIArgProSerTyrlasn.. 26
318 CACCTCTGTCTGTCTGATGACATTAGCCAGGAGGAGAGAGAGCGCTGGCG 367
      ::::::::::: ||| ::::::::::: |||
27 .....IleAlaGluAsp.....ThrThrLeuGluP 35
368 CCGAC.....AGTCTCTGTGGGGGTCAAGAGAC.....CCCGCAGC 405
      ::::::::::: ||| ::::::::::: |||
35 roGluAspAlaAsnSerSerGlyAspAspGluAspAspAspAspGlySer 51
406 CAGCAGGGGCGACGA.....CCGCGCTTCACACA 434
      ::::::::::: ||| ::::::::::: |||
52 GluAspPheThrAsnAspAsnAsnHisMetArgAlaProTyrlThrPnas 68
435 GCCCTCAAGATGAGCGCGCGGTGATCGCAGCGCGCGCTGAGTGTCCG 484
      ::::::::::: ||| ::::::::::: |||
68 nThrGluIySleuGluIySlySleuHisAlaValProAlaAlaAsnThrV 85
485 TGGCGGTCAAGTGTGGCCAGCGCGCGCGCGCGCGCGCGCGCGCATCAGTGG 534
      ::::::::::: ||| ::::::::::: |||
85 allySphaArgCysProAlaGlyGlyAsnProThrProSerMetArgTyrlP 101
535 ATGAAGAGACACACAGGCGCTTGAGC.....CGCCACAGGCGCGGTGA 575
      ::::::::::: ||| ::::::::::: |||
102 LeuIyAsnGlyIySgluPheIySgluGlnIySArgIleGlyIyPheIy 118
576 GCCCAGAGAGAGAGAGATGAGACTGAGCCTGAGAACCTGCGCGCGAGG 625
      ::::::::::: ||| ::::::::::: |||
118 sValArgSerGlnHisPheSerLeuIleMetGluSerValValProSerA 135
626 ACAGCGGCAAAATACCTCGCGGTGTGCAACCGCGCGCGCGCGCATAC 675
      ::::::::::: ||| ::::::::::: |||
135 spGluIyAsnTyrlThncysIleMetGluAsnGluTyrlGlySerIleAsn 151
676 GCCACTACAGAGTGTGATGATCCAGCGGACCGCGTTCACAGCCCGTGT 725
      ::::::::::: ||| ::::::::::: |||
152 HisThrTyrlHisLeuAspValAlaGluArgSerProHisArgProIle 168
726 CACAGGACGACCGCGGTGAACAGAGCGGTGAGCTTGGGGGAGCAGCGT 775
      ::::::::::: ||| ::::::::::: |||
168 uGlnAlaIyGluProAlaAsnThrThrIySValGlyIyAspAlaG 185
776 CTTTCAGTGTCAAGGTGTGCGCAGCGAGCTGAAGCGGTGATCCAGTGTG 825
      ::::::::::: ||| ::::::::::: |||
185 luhValAlcylslyValTyrlSerAspAlaGlnProHisIleGlnTyrlPle 201
826 AACCGGTGTGAGTACGGCGCGGAGGCGCGGACACATCCACATCATGT 875
      ::::::::::: ||| ::::::::::: |||
202 ArgHisPheGluLeu.....AsnGlySerIySleGlyP 213
876 GGGCGGCGCAGAGTTGTG...GTGCTGCCAGGAGGAGTGTGTGTGCG 922
      ::::::::::: ||| ::::::::::: |||
213 oAspGlyHisProTyrlLeuIySValLeuIySArgSerGlyIle..... 227
923 GGGCGGAGCGTCTACATTAATAGCTGTACACCGTGTGCGCGCGAG 972
      ::::::::::: ||| ::::::::::: |||
228 .....AsnSerSerAsnAlaGluValLeuThrLeuHisAsnValThnGlu 242
973 GACGATCGGGCATGTACATCTGCTTGGGCGGCAACACATGGGCGTACAG 1022
      ||| ::::::::::: ||| ::::::::::: |||

```

```
243 AlaasparglyIntYrThrcysValserasnYrIleGlyIuAl 259
1023 CTTCCGAGCGCCTTCTCAGCCGTGCGCA..GACCCAAACCGCCAG 1069
      : :::::::::::::::::::::::::::: :
259 aasnGlnSerAlatrPleuThrValleuProAlaSerGluYaspspGluG 276
      : :::::::::::::::::::: :
1070 GGCACACTGTGGCTCTGCTGCTGCGCACTAGCTGCGCGCGGTG 1119
      : :::::::::::::::::::: :
276 luArgGluLeuAspSerSerSerGluYrThrGluIle..... 288
1120 GTCATCGCATCCAGCGCGCTGCTTCATCTGGGACCGCTGCTCT 1169
      : :::::::::::::::::::: :
289 .....AlaIleYrCysValGlyGlyPheIuI 298
1170 G 1170
298 e 298

seq_name: /cgn2_6/prodata/2/laa/5B_COMB.pep:US-08-885-418-10

seq documentation block:
; Sequence ID, Application US/08885418
; Patent No. 5923528
; GENERAL INFORMATION:
; APPLICANT: Chiu, Ing-Ming
; APPLICANT: Poulin, Matthew L
; TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ing-Ming Chiu
; STREET: S2052 Davis Medical Research Center, 480 West
; CITY: Columbus
; STATE: Ohio
; COUNTRY: USA
; ZIP: 43210
; COMPUTER-READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/885,418
; FILING DATE:
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (614)-293-8093
; TELEFAX: (614)-293-5631
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 731 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-885-418-10

alignment_scores:
      Quality: 339.50      Length: 317
      Ratio: 1.741      Gaps: 11
      Percent Similarity: 61.514      Percent Identity: 29.338

alignment block:
US-09-598-042a-2 x US-08-885-418-10 ..

Align seg 1/1 to: US-08-885-418-10 from: 1 to: 731

268 GTTAGCGTGTCAAGCCACCAAGCGCTGGCAGCGCTGACGTCACATA 317
      ||| ||||| :||| :||| :|||
11 ValmetValAlatrAlatrPleuSerleuAlaArgProSerTYrAsn.. 26
      : :::::::::::::::::::: :
318 CACCTGCTGCTGCTGATGACATTAGCCAGGAGAGAGACCTGGGCG 367
      : :::::::::::::::::::: :
```

```
27 .....IleAlaGluAsp.....ThrThrleuGluP 35
368 CCGAGC.....AGCTCTGTGGGGGTCAAGAGAC.....CCGGCCAGC 405
      : :::::::::::::::::::: :
35 roGluAspAlaAsnSerSerleYaspspGluAspAsnAspGlySer 51
      : :::::::::::::::::::: :
406 CAGCAGTGGGACAGA.....CCGGCTTCACACA 434
      : :::::::::::::::::::: :
52 GluAspPheThrAsnAspAsnHisMetArgAlaProYrTrPThrAs 68
      : :::::::::::::::::::: :
435 GCGCTCCAGATGAGCGCGGTGATGCGACGCGCGGTGGTAGTCCG 484
      : :::::::::::::::::::: :
68 ThrGluIuSerleuGluYrCysLeuHisAlaValProAlaIaAsnThr 85
485 TGGCGGTCAAGTGGTGGCCAGCGGCGACCTCGCGCGCATACGTTG 534
      : :::::::::::::::::::: :
85 allYsPheArgCysProAlaGlyIuYasProThrProSerMetArgTrp 101
535 ATGAGAGACGACAGCGCTTGACG.....CGCCAGAGCGCGGTGA 575
      : :::::::::::::::::::: :
102 leuYasasnGlyYsGluPheYsGlnGluHisArgIleGlyGlyPheY 118
576 GCCCAGGAGAGAGAGTGGACACTGAGCGCTGAGAACCTGCGCGCGAG 625
      : :::::::::::::::::::: :
118 sValArgSerGlnHisPheSerleuIleMetGluSerValValProSer 135
626 ACAGCGGCAATACACCTCGCGCGGTGCGAGCGCGGCGCGCATCAC 675
      : :::::::::::::::::::: :
135 spGluGlyAsnYrThrcysIleMetGluAsnGluYrGlySerIleAsn 151
676 GCCACACTACAGGTGATGTGATGCAGCGGACCGCTTCACAGCGCGTCT 725
      : :::::::::::::::::::: :
152 HisThrYrHisleuAspValValGluArgSerProHisArgProIle 168
726 CACAGGACAGCAGCCCGTGAACACAGAGGTGACTTGGGGGACACAGT 775
      : :::::::::::::::::::: :
168 uGluAlaGlyLeuProAlaAsnThrThrThrLysValGlyIuYasPhe 185
776 CTTCCAGTGCAGAGTGGCGGAGCGAGCGGAGCGGTGATCAGTGGCTG 825
      : :::::::::::::::::::: :
185 IuPheValCysLysValYrSerAspAlaGlnProHisIleGlnTrpIle 201
826 AAGCGCGTGAAGTACGGCGCGGAGCGCGCCACCAACTCCACATCGATG 875
      : :::::::::::::::::::: :
202 ArgHisPheGluLeu.....AsnGlySerLysIleGlyPyr 213
876 GGGCGGCCAGAGTTTGTG...GTGCTGCCAGCGGTGACGTGTGTGCG 922
      : :::::::::::::::::::: :
213 aAspGlyHisProYrleuYsValleuYsArgSerleYrIle..... 227
923 GCGCCGAGCGGCTCTACCTCAATAGCTGTCTATCACCAGCGCGCGCAG 972
      : :::::::::::::::::::: :
228 .....AsnSerSerAsnAlaGluValleuThrleuHisAsnValThrGlu 242
973 GACGATGGGCGCATGTACATCTGCTTGGCGCAACACACATGGCTACAG 1022
      : :::::::::::::::::::: :
243 AlaasparglyIntYrThrcysLysValserasnYrIleGlyIuAl 259
1023 CTTCCGAGCGCCTTCTCAGCCGTGCGCA..GACCCAAACCGCCAG 1069
      : :::::::::::::::::::: :
259 aasnGlnSerAlatrPleuThrValleuProAlaSerGluYaspspGluG 276
1070 GGCACACTGTGGCTCTGCTGCTGCGCACTAGCTGCGCGCGGTG 1119
      : :::::::::::::::::::: :
276 luArgGluLeuAspSerSerSerGluYrThrGluIle..... 288
1120 GTCATCGCATCCAGCGCGCTGCTTCATCTGGGACCGCTGCTCTCT 1169
      : :::::::::::::::::::: :
289 .....AlaIleYrCysValGlyGlyPheIuI 298
1170 G 1170
298 e 298
```



seq\_name: /cgn2\_6/ptodata/2/1ae/5A\_COMB.pep:US-07-640-029-5

seq\_documentation\_block:

Sequence 5, Application US/07640029

Patent No. 5229501

GENERAL INFORMATION:

APPLICANT: Kiefer, Michael C.

APPLICANT: Valenzuela, Pablo D.T.

APPLICANT: Barr, Phillip J.

TITLE OF INVENTION: Expression and Use of Human Fibroblast

NUMBER OF INVENTION: Growth Factor Receptor

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation

STREET: 4560 Horton Street

CITY: Emeryville

STATE: California

COUNTRY: USA

ZIP: 94608

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/640,029

FILING DATE: 19910111

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: McClung, Barbara G.

REGISTRATION NUMBER: 33,113

REFERENCE/DOCKET NUMBER: CH-165

TELECOMMUNICATION INFORMATION:

TELEPHONE: 510-655-3542

TELEFAX: 510-601-2708

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 300 amino acids

TYPE: AMINO ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-07-640-029-5

alignment\_scores:

Quality: 308.00

Ratio: 2.425

Percent Similarity: 72.989

Percent Identity: 35.632

Length: 174

Gaps: 3

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

Percent Identity: 35.632

seq\_name: /cgn2\_6/ptodata/2/1ae/6B\_COMB.pep:US-08-439-992A-5

seq\_documentation\_block:

Sequence 5, Application US/08439992A

Patent No. 625454

GENERAL INFORMATION:

APPLICANT: Kiefer, Michael C.

APPLICANT: Pablo, Valenzuela D.T.

APPLICANT: Phillip, Barr J.

TITLE OF INVENTION: Expression and Use of Human Fibroblast

NUMBER OF INVENTION: Receptor

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation

STREET: 4560 Horton Street

CITY: Emeryville

STATE: CA

COUNTRY: USA

ZIP: 94608

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/439,992A

FILING DATE: 12-MAY-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Chung, Ling-fong

REGISTRATION NUMBER: 36,482

REFERENCE/DOCKET NUMBER: 0165,004

TELECOMMUNICATION INFORMATION:

TELEPHONE: 510-923-2704

TELEFAX: (510) 655-3542

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 300 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-439-992A-5

alignment\_scores:



795 CAGCGACGTGAAGCCGGTGTATCCACTGGCTGAGCGCGTGAG 837  
||||| :||| |||||||||:|||||  
191 rserasproginprohislleglntprleuylshstlleglu 205  
seq\_name: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep:US-08-441-944A-7  
seq\_documentation\_block:  
: Sequence 7, Application US/08441944A  
: Patent No. 5767250  
: GENERAL INFORMATION:  
: APPLICANT: SPAETE, RICHARD  
: TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION  
: NUMBER OF SEQUENCES: 20  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: CHIRON CORPORATION  
: STREET: 4560 Horton Street - R440  
: CITY: Emeryville  
: STATE: CA  
: COUNTRY: USA  
: ZIP: 94608-2916  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: OPERATING SYSTEM: IBM PC compatible  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/441,944A  
: FILING DATE: 16-MAY-1995  
: CLASSIFICATION: 530  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 07/921,807  
: FILING DATE: 29-SEP-1992  
: ATTORNEY/AGENT INFORMATION:  
: NAME: MCCLUNG, BARBARA G.  
: REGISTRATION NUMBER: 33,113  
: REFERENCE/DOCKET NUMBER: 0209,001  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (510) 601-2708  
: TELEFAX: (510) 655-3542  
: INFORMATION FOR SEQ ID NO: 7:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 302 amino acids  
: TYPE: amino acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
: US-08-441-944A-7  
alignment\_scores:  
Quality: 306.00 Length: 181  
Ratio: 2.448 Gaps: 4  
Percent Similarity: 69.061 Percent Identity: 34.254  
alignment\_block:  
US-09-598-042A-2 x US-08-441-944A-7 ..  
Align seg 1/1 to: US-08-441-944A-7 from: 1 to: 302  
337 GACATTAGCCCGAGAGAGAGCGCTGGGCGCCGACAGCTCCTCTGGGGG 386  
||||| :||| |||||||||:|||||  
31 aspalaleuprosertergluspaspaspaspaspaspaserse 47  
387 TCAAGAG.....GACCCCGCCAGCC 406  
:||||| :||| |||||||||:|||||  
47 rclululysglutlrpaspasrthlyspasnpaspasnpasnpasnpas 64  
407 AGCACTGGGCGACCGCGCTTACACACAGCCCTCCAAATGAGGCGCGG 456  
||||| :||| |||||||||:|||||  
64 rotlytrp.....ThsSerProglulysmetglulyslys 75  
457 GTGATCGCAGCGCCCGTGGGTAGCTCGTGGCGCTCAAGTGGTGGCCAG 506

76 LeuHsAlaValProAlaAlaLysThrValLysPheLysCysProserse 92  
:|:|:| :||| :|||:|||||:|||||:||||| :|||  
507 CGGCGACCCCTCGCCGACATCACGTGATGAGAGACGACGACGCTTGA 556  
||||| :|||:||||| :|||:|||||:||||| :|||  
92 rcllythrproasnprothlrleuargtrpseuylsasnlglyslglulys 108  
557 CGGCGCCAGAG.....GCCGCTGAGCCGACGAGAGAGAGTGG 594  
:||||| :||| :||| :|||  
109 ..LysProasphtlsarglleglygltyrlyslvalargtyralatrrtp 124  
595 ACACGTAGCCCTGAAGACCTGGCGCGGAGACGAGCGGCAATTAACCTG 644  
:||||| :||| :||| :|||  
125 Serllellemetaspservalalproseraspysgllyasrlythrty 141  
645 CCGCGTGTGACACCGCGCGCGCCATCAACGCCACTTCAGAGTGGATG 694  
| :|||:||||| :|||:||||| :|||:||||| :|||  
141 sllevalgluasnglultrygrlyserlleasnhtsrthrygrlnleuaspv 158  
695 TGATCCAGCGGACCCGTTCCAAAGCCGCTGTCAACAGGACGACCCCGTG 744  
|:|||||:||||| :|||:||||| :|||:||||| :|||  
158 alvalgluargserprohlsarproilleuaglalaaglyleuproala 174  
745 AACAGAGCGGTGACCTGGGGGAGCACGCTCTCCAGTGCAGAGTGGG 794  
||| :||| :|||:|||||:|||||:|||||:|||||  
175 AsnLysThrValAlaLeuGlyserhsnvalgluphemecyslyslvalty 191  
795 CAGCGACGTGAAGCCGGTGTATCCACTGGCTGAGCGCGTGAG 837  
||||| :||| |||||||||:|||||  
191 rserasproginprohislleglntprleuylshstlleglu 205  
seq\_name: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep:US-07-640-029-6  
seq\_documentation\_block:  
: Sequence 6, Application US/07640029  
: Patent No. 5229501  
: GENERAL INFORMATION:  
: APPLICANT: Kiefer, Michael C.  
: APPLICANT: Valenzuela, Pablo D.T.  
: TITLE OF INVENTION: Expression and Use of Human Fibroblast  
: NUMBER OF SEQUENCES: 12  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Chiron Corporation  
: STREET: 4560 Horton Street  
: CITY: Emeryville  
: STATE: California  
: COUNTRY: USA  
: ZIP: 94608  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: OPERATING SYSTEM: IBM PC compatible  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/07/640,029  
: FILING DATE: 19910111  
: CLASSIFICATION: 530  
: ATTORNEY/AGENT INFORMATION:  
: NAME: MCCLUNG, BARBARA G.  
: REGISTRATION NUMBER: 33,113  
: REFERENCE/DOCKET NUMBER: CH-165  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 510-601-2708  
: TELEFAX: 510-655-3542  
: INFORMATION FOR SEQ ID NO: 6:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 302 amino acids  
: TYPE: AMINO ACID  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: peptide



760 TTCGGGGGACCCAGCTCTTCAGTCAGGTGCGGACGCGTGAAGCC 809  
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
180 LeuGlySerAsnValGluPheMetCysLysValTyrSerAspProGlnP 196  
810 GGTGATCCAGTGGCTGAGACGGCGGTGAG 837  
| |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
196 ohStIleGlnTPrLeuLysHisIleGlu 205

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-08-441-944A-8

seq\_documentation\_block:

; Sequence 8, Application US/08441944A  
; Patent No. 5767250  
; GENERAL INFORMATION:  
; APPLICANT: SPAETE, RICHARD  
; TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: CHIRON CORPORATION  
; STREET: 4560 Horton Street - R440  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/441,944A  
; FILING DATE: 16-MAY-1995  
; CLASSIFICATION: 530  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 07/921,807  
; FILING DATE: 29-SEP-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MCCLEUNG, BARBARA G.  
; REGISTRATION NUMBER: 33,113  
; REFERENCE/DOCKET NUMBER: 0209.001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 601-2708  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 302 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-441-944A-8

alignment\_scores:

Quality: 304.00 Length: 176  
Ratio: 2.394 Gaps: 3  
Percent Similarity: 72.159 Percent Identity: 34.659

alignment\_block:

US-09-598-042A-2 x US-08-441-944A-8 ..

Align seg 1/1 to: US-08-441-944A-8 from: 1 to: 302

337 GACATATGACCCAGGAGAGAGACCTGGGGCCGACAGCTCTCTGGGG 386  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
31 AspaLalelProSerSerGluAspAspAspAspAspAspAspAsp 47  
387 TCAGAGAGACCCCGCC.....AGCCAGAGTGGGACGAC 421  
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||  
47 rgluglulysglulthrAspAsnThrLysProAsnArgMetProValAla 64  
422 CGGCTTCAACAGACCCCTCCAGATGAGCGCGGGTGTATCGACGCGCC 471

|| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
64 roTyrTrpHisSerProGluLysMetGluLysLysLeuHisAlaValPro 80  
472 GTGGTAGCTCCGTGGCGCTCAAGTCCGTGGCCAGCGGCGACCCCTGGCC 521  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
81 AlaAlaLysThrValLysPheLysCysProSerSerGlyThrProAsnPr 97  
522 CGACATCAGCTGATGAGGAGCCAGCCAGCTGAGCGCGCCAG..... 567  
| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
97 oThrLeuArgTPrLeuLysAsnLysGluPhe...LysProAsnHisA 113  
568 .....GCCGCTGAGCCCGAGAGAGAGAGTGAAGTGAAGTGAAG 609  
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||  
113 rglleGlyGlyTyrLysValArgTyrAlaThrTrpSerIleIleMetAsp 129  
610 AACCTGCGCGCGAGAGACAGCGCAAAATACCTGCGCGGTGTCGACCG 659  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
130 SerValValProSerAspLysGlyAsnTyrThrCysIleValGluAsnGl 146  
660 CGGGGCGCCATCAAGCCACTCAAGTGTGATGATCCAGCGGACCC 709  
: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
146 uTyrGlySerIleAsnHisThrTyrGlnLeuAspValValGluArgSerP 163  
710 GTTCCAGCCCGTGTCAACAGCCAGCCCGCTGAACAGAGCTGGAC 759  
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||  
163 roHisArgProIleLeuGlnAlaGlyLeuProAlaAsnLysThrValAla 179  
760 TTCGGGGGACCCAGCTCTTCAGTCAGGTGCGGACGCGTGAAGCC 809  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
180 LeuGlySerAsnValGluPheMetCysLysValTyrSerAspProGlnP 196  
810 GGTGATCCAGTGGCTGAGACGGCGGTGAG 837  
| |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
196 ohStIleGlnTPrLeuLysHisIleGlu 205

seq\_name: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:US-08-439-992A-6

seq\_documentation\_block:  
; Sequence 6, Application US/08439992A  
; Patent No. 6255454  
; GENERAL INFORMATION:  
; APPLICANT: Kiefer, Michael C.  
; APPLICANT: Pablo, Valenzuela D.T.  
; APPLICANT: Philip, Barr J.  
; TITLE OF INVENTION: Expression and Use of Human Fibroblast  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/439,992A  
; FILING DATE: 12-MAY-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Chung, Ling-Fong  
; REGISTRATION NUMBER: 36,482  
; REFERENCE/DOCKET NUMBER: 0165.004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 510-923-2704  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 302 amino acids

```

;      TYPE: amino acid
;      STRANDEDNESS: single
;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
US-08-439-992A-6

```

```
alignment_scores:
  Quality: 304.00      Length: 176
  Ratio: 2.394      Gaps: 3
Percent Similarity: 72.159      Percent Identity: 34.659
```

```
alignment_block:
```

Align seg 1/1 to: US-08-439-992A-6 from: 1 to: 302

[illegible]



PA (HUMA-) HUMAN GENOME SCI INC.  
 XX Ruben SM, Young PE;  
 XX WPI: 2000-387035/33.  
 DR P-PSDB: AAY92864.  
 XX  
 XX Nucleic acids encoding fibroblast growth factor-5 useful for the  
 PT prevention, diagnosis and treatment of conditions associated with  
 PT tissue repair and aberrant cell functions, e.g. cell survival and  
 PT proliferation  
 PS Claim 1: Fig 1A-C; 182pp; English.  
 XX  
 XX This DNA encodes a fibroblast growth factor receptor protein, FGFR-5. The  
 CC FGFR-5 protein and DNA may be used in the prevention, treatment and  
 CC diagnosis of diseases or conditions associated with inappropriate FGFR-5  
 CC expression and activity. For example, the nucleic acids (and vectors  
 CC containing them) and the FGFR-5 polypeptide may be used to treat  
 CC disorders associated with increased or decreased cell survival (such as  
 CC cancers (e.g. leukemia, colonic cancer, testicular cancer and follicular  
 CC lymphomas), autoimmune disorders (e.g. multiple sclerosis and Crohn's  
 CC disease) viral infections (e.g. herpes viruses), inflammation, graft  
 CC versus host disease, acute and chronic graft rejection, ischemic injuries  
 CC and atherosclerosis), activation, secretion, migration, differentiation  
 CC and proliferation, diseases associated with defects in wound healing,  
 CC mucopolys, defects of angiogenesis, immune dysfunction, endocrine  
 CC dysfunction and insulin secretion disorders. Anti-FGFR-5 antibodies may  
 CC also be used as diagnostic agents for detecting the presence of FGFR-5  
 CC polypeptides in samples.  
 XX  
 XX Sequence 3112 BP; 686 A; 1035 C; 878 G; 513 T; 0 other;  
 SQ

Query Match 88.3%; Score 2814.4; DB 21; Length 3112;  
 Best Local Similarity 95.5%; Pred. No. 0; Mismatches 11; Indels 128; Gaps 1;  
 Matches 2259; Conservative 0;

QY 89 ttcgttcgcaagtgccggtcgtggtccgacatctgcgtcagtcgccaagtggaggg 148  
 Db 131 tcccaagcagagtgccggtcgtggtccgacatctgcgtcagtcgccaagtggaggg 190  
 QY 149 accgcgcgcgtcgtcagatctgacaaagatggtccgcacatccacagcgttggagcc 208  
 Db 191 accgcgcgcgtcgtcagatctgacaaagatggtccgcacatccacagcgttggagcc 250  
 QY 209 gcttcgcgctgtcgtcgcagcgggtcgtgaagtgaaagcgtgagggaggtgctgcggcg 268  
 Db 251 gcttcgcgctgtcgtcgcagcgggtcgtgaagtgaaagcgtgagggaggtgctgcggcg 310  
 QY 269 tctacgtgtgcaagcgcacaaagcgttcgcgacgcttgagcgtlcaactacacccctcgtc 328  
 Db 311 tctacgtgtgcaagcgcacaaagcgttcgcgacgcttgagcgtlcaactacacccctcgtc 370  
 QY 329 tctgtgatagactttagccagggagagagccttgggcccgcacagctcctctgggggtc 388  
 Db 371 tctgtgatagactttagccagggagagagccttgggcccgcacagctcctctgggggtc 430  
 QY 389 aagaggaaccccgacagcagctggtgcaagcgcgtcttccacagccctcccaagatga 448  
 Db 431 aagaggaaccccgacagcagctggtgcaagcgcgtcttccacagccctcccaagatga 490  
 QY 449 ggcgcgcggtgctgcacagcgcgttgggtagctccgtgcgtcctcaagtgtcgtgcacag 508  
 Db 491 ggcgcgcggtgctgcacagcgcgttgggtagctccgtgcgtcctcaagtgtcgtgcacag 550  
 QY 509 ggcacccctcgcgcgcacatcaagtgtgataagagagcagcagccttgcgcgcgcagag 568  
 Db 551 ggcacccctcgcgcgcacatcaagtgtgataagagagcagcagccttgcgcgcgcagag 610  
 QY 569 ccgcgtgaccccaagagaagatgagacatcagcttgaagaaccttggcgcgagagaca 628  
 Db 611 ccgcgtgaccccaagagaagatgagacatcagcttgaagaaccttggcgcgagagaca 670

QY 629 ggcggcaatatcacctgcccgtgttcgaaccgcggcgccatcaacgccaactacaag 688  
 Db 671 ggcggcaatatcacctgcccgtgttcgaaccgcggcgccatcaacgccaactacaag 730  
 QY 689 tgcattgtatccagcagcagccttccaaagccgtgtcctacagcagcagcagcagcagc 748  
 Db 731 tgcattgtatccagcagcagccttccaaagccgtgtcctacagcagcagcagcagcagc 790  
 QY 749 cgaaggtgacttcgggggagacacagctcctccagtgcaagtgccgagcagctgaagc 808  
 Db 791 cgaaggtgacttcgggggagacacagctcctccagtgcaagtgccgagcagctgaagc 850  
 QY 809 cgttatccagtggtcgtgaagcgttggagtgagcgcgcgagggccgcacacactccaca 868  
 Db 851 cgttatccagtggtcgtgaagcgttggagtgagcgcgcgagggccgcacacactccaca 910  
 QY 869 tgcattgtggtgcgcgcagaaagtgtgtgtcgtccacagcagtgacgtgtgtgcgcgcgc 928  
 Db 911 tgcattgtggtgcgcgcagaaagtgtgtgtcgtccacagcagtgacgtgtgtgcgcgcgc 970  
 QY 929 acggtctcctacatcaataagctgtcctacacccgtgcgcgcgcagcagatgcggcagt 988  
 Db 971 acggtctcctacatcaataagctgtcctacacccgtgcgcgcgcagcagatgcggcagt 1030  
 QY 989 acatcgtccttgccgcacacacatggggtacagcttccgcgcgcgccttcacacgttc 1048  
 Db 1031 acatcgtccttgccgcacacacatggggtacagcttccgcgcgcgccttcacacgttc 1090  
 QY 1049 tgcagaaacccaaacgcgcagcagcagctgtgctcctcgtcctcgcgcagcagcagc 1108  
 Db 1091 tgcagaaacccaaacgcgcagcagcagctgtgctcctcgtcctcgcgcagcagcagc 1150  
 QY 1109 cgtggtccgtgtgtatcgtgcatccagcgcgcgtcttctacactctgtggcagcctgtcc 1168  
 Db 1151 cgtggtccgtgtgtatcgtgcatccagcgcgcgtcttctacactctgtggcagcctgtcc 1210  
 QY 1169 tctggtcttgccagcgcgcagaaagacgctgtgacccgcgcgcgtcctcctcgtcgtc 1228  
 Db 1211 tctggtcttgccagcgcgcagaaagacgctgtgacccgcgcgcgtcctcctcgtcgtc 1270  
 QY 1229 ggcacgc 1288  
 Db 1271 ggcacgc 1330  
 QY 1289 ccgcctcagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1348  
 Db 1331 ccgcctcagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1359  
 QY 1349 ccagcacttactggtccaggtccagctgtcgtccctaaagtgtaccccaactctaca 1408  
 Db 1360 ----- 1359  
 QY 1409 cagacatccacaac 1468  
 Db 1360 ----- -ctgtgtgagagacatggctcc 1382  
 QY 1469 ggcagcccccagcacttactggtggccagggccaggttctgtgcccctaagtgtacccaa 1528  
 Db 1383 ggcagcccccagcacttactggtggccagggccaggttctgtgcccctaagtgtacccaa 1442  
 QY 1529 actctacacagacatccacaac 1588  
 Db 1443 actctacacagacatccacaac 1502  
 QY 1589 ggcacaggttccacacagac 1648  
 Db 1503 ggcacaggttccacacagac 1562  
 QY 1649 acggtgggggc 1708  
 Db 1563 acggtgggggc 1622





Query Match	87.9%;	Score 2801.6;	DB 22;	Length 3169;
Best Local Similarity	95.3%;	Pred. No. 0;		
Matches 2951;	Conservative	0;	Mismatches 19;	Indels 128;
			Gaps	1

[illegible]



[illegible][illegible]

[illegible]

Db	3306	ccccaccaccatgctgctgggccccagatctctgtaatttattgttaagatttgagctga	3365
Oy	3122	agccccgtatatattacatttttgttaacaatga	3158
Db	3366	agccccgtatatattacatttttgttaacaacaaa	3402
RESULT 4			
ID	AAZ64984		
XX	AAZ64984 standard; cDNA; 3402 BP.		
XX	AAZ64984;		
XX			
DT	05-APR-2000 (first entry)		
DE	Membrane-bound protein PRO943 encoding cDNA.		
XX			
KW	Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;		
KW	pharmaceutical; receptor immunoadhesin; gene mapping; ss.		
OS	Homo sapiens.		
XX			
PM	W09963088-A2.		
XX			
PD	09-DEC-1999.		
XX			
PF	02-JUN-1999; 99WO-US12252.		
XX			
PR	02-JUN-1998; 98US-0087607.		
PR	02-JUN-1998; 98US-007609.		
PR	02-JUN-1998; 98US-0087159.		
PR	03-JUN-1998; 98US-0087827.		
PR	04-JUN-1998; 98US-0088021.		
PR	04-JUN-1998; 98US-0088025.		
PR	04-JUN-1998; 98US-0088028.		
PR	04-JUN-1998; 98US-0088029.		
PR	04-JUN-1998; 98US-0088030.		
PR	04-JUN-1998; 98US-0088033.		
PR	04-JUN-1998; 98US-0088326.		
PR	05-JUN-1998; 98US-0088167.		
PR	05-JUN-1998; 98US-0088202.		
PR	05-JUN-1998; 98US-0088212.		
PR	05-JUN-1998; 98US-0088217.		
PR	09-JUN-1998; 98US-0088655.		
PR	10-JUN-1998; 98US-0088722.		
PR	10-JUN-1998; 98US-0088730.		
PR	10-JUN-1998; 98US-0088734.		
PR	10-JUN-1998; 98US-0088738.		
PR	10-JUN-1998; 98US-0088740.		
PR	10-JUN-1998; 98US-0088741.		
PR	10-JUN-1998; 98US-0088742.		
PR	10-JUN-1998; 98US-0088810.		
PR	10-JUN-1998; 98US-0088811.		
PR	10-JUN-1998; 98US-0088824.		
PR	10-JUN-1998; 98US-0088825.		
PR	10-JUN-1998; 98US-0088826.		
PR	11-JUN-1998; 98US-0088858.		
PR	11-JUN-1998; 98US-0088861.		
PR	11-JUN-1998; 98US-0088863.		
PR	11-JUN-1998; 98US-0088876.		
PR	12-JUN-1998; 98US-0089090.		
PR	12-JUN-1998; 98US-0089105.		
PR	16-JUN-1998; 98US-0089440.		
PR	16-JUN-1998; 98US-0089512.		
PR	16-JUN-1998; 98US-0089514.		
PR	17-JUN-1998; 98US-0089532.		
PR	17-JUN-1998; 98US-0089538.		
PR	17-JUN-1998; 98US-0089598.		
PR	17-JUN-1998; 98US-0089599.		
PR	17-JUN-1998; 98US-0089600.		
PR	17-JUN-1998; 98US-0089653.		
PR	18-JUN-1998; 98US-0089801.		

```

PR 18-JUN-1998; 98US-0089907.
PR 18-JUN-1998; 98US-0089908.
PR 19-JUN-1998; 98US-0089947.
PR 19-JUN-1998; 98US-0089948.
PR 19-JUN-1998; 98US-0089952.
PR 22-JUN-1998; 98US-0090246.
PR 22-JUN-1998; 98US-0090252.
PR 22-JUN-1998; 98US-0090254.
PR 23-JUN-1998; 98US-0090349.
PR 23-JUN-1998; 98US-0090355.
PR 24-JUN-1998; 98US-0090429.
PR 24-JUN-1998; 98US-0090431.
PR 24-JUN-1998; 98US-0090435.
PR 24-JUN-1998; 98US-0090444.
PR 24-JUN-1998; 98US-0090445.
PR 24-JUN-1998; 98US-0090472.
PR 24-JUN-1998; 98US-0090533.
PR 24-JUN-1998; 98US-0090538.
PR 24-JUN-1998; 98US-0090540.
PR 24-JUN-1998; 98US-0090557.
PR 25-JUN-1998; 98US-0090676.
PR 25-JUN-1998; 98US-0090678.
PR 25-JUN-1998; 98US-0090688.
PR 25-JUN-1998; 98US-0090690.
PR 25-JUN-1998; 98US-0090691.
PR 25-JUN-1998; 98US-0090694.
PR 25-JUN-1998; 98US-0090695.
PR 25-JUN-1998; 98US-0090696.
PR 26-JUN-1998; 98US-0090862.
PR 26-JUN-1998; 98US-0090863.
PR 01-JUL-1998; 98US-0091358.
PR 01-JUL-1998; 98US-0091544.
PR 02-JUL-1998; 98US-0091478.
PR 02-JUL-1998; 98US-0091486.
PR 02-JUL-1998; 98US-0091519.
PR 02-JUL-1998; 98US-0091626.
PR 02-JUL-1998; 98US-0091628.
PR 02-JUL-1998; 98US-0091633.
PR 02-JUL-1998; 98US-0091646.
PR 02-JUL-1998; 98US-0091673.
PR 07-JUL-1998; 98US-0091978.
PR 09-JUL-1998; 98US-0091982.
PR 10-JUL-1998; 98US-0092472.
PR 20-JUL-1998; 98US-0093339.
PR 30-JUL-1998; 98US-0094651.
PR 04-AUG-1998; 98US-0095282.
PR 04-AUG-1998; 98US-0095285.
PR 04-AUG-1998; 98US-0095301.
PR 04-AUG-1998; 98US-0095302.
PR 04-AUG-1998; 98US-0095318.
PR 04-AUG-1998; 98US-0095321.
PR 04-AUG-1998; 98US-0095325.
PR 10-AUG-1998; 98US-0095916.
PR 10-AUG-1998; 98US-0095929.
PR 11-AUG-1998; 98US-0096012.
PR 11-AUG-1998; 98US-0096143.
PR 11-AUG-1998; 98US-0096146.
PR 12-AUG-1998; 98US-0096329.
PR 17-AUG-1998; 98US-0096757.
PR 17-AUG-1998; 98US-0096766.
PR 17-AUG-1998; 98US-0096768.
PR 17-AUG-1998; 98US-0096773.
PR 17-AUG-1998; 98US-0096791.
PR 17-AUG-1998; 98US-0096867.
PR 17-AUG-1998; 98US-0096891.
PR 17-AUG-1998; 98US-0096894.
PR 17-AUG-1998; 98US-0096895.
PR 17-AUG-1998; 98US-0096897.
PR 18-AUG-1998; 98US-0096949.
PR 18-AUG-1998; 98US-0096950.

```

```

PR 18-AUG-1998; 98US-0096959.
PR 18-AUG-1998; 98US-0096960.
PR 18-AUG-1998; 98US-0096962.
PR 19-AUG-1998; 98US-0097141.
PR 20-AUG-1998; 98US-0097218.
PR 24-AUG-1998; 98US-0097661.
PR 26-AUG-1998; 98US-0097951.
PR 26-AUG-1998; 98US-0097952.
PR 26-AUG-1998; 98US-0097954.
PR 26-AUG-1998; 98US-0097955.
PR 26-AUG-1998; 98US-0097971.
PR 26-AUG-1998; 98US-0097974.
PR 26-AUG-1998; 98US-0097978.
PR 26-AUG-1998; 98US-0097979.
PR 26-AUG-1998; 98US-0097986.
PR 26-AUG-1998; 98US-0098014.
PR 31-AUG-1998; 98US-0098025.
PR 16-SEP-1998; 98US-0100634.
PR 12-JAN-1999; 99US-0115565.

```

(GETH ) GENENTECH INC.

Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;  
Wood WI, Yuan J;

WPI, 2000-072883/06.

P-PSDB; AAY66656.

Membrane-bound proteins and related nucleotide sequences

Claim 2; Fig 69; 822pp; English.

The invention provides membrane-bound PRO polypeptides and polynucleotides encoding them. The PRO sequences of the invention were identified based on extracellular domain homology screening. The PRO sequences have homology with proteins including LDL receptors, TIE CC ligands and various enzymes. The membrane-bound proteins and receptor CC molecules are useful as pharmaceutical and diagnostic agents. Receptor CC immunoadhesins, for instance, can be used as therapeutic agents to block receptor-ligand interactions. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. The PRO encoding sequences are useful as hybridization probes, in chromosome and gene mapping and in the generation of antisense RNA and DNA. PRO nucleic acid sequences will also be useful for the preparation of PRO polypeptides, especially by recombinant techniques.

Sequence 3402 BP; 730 A; 1151 C; 976 G; 545 T; 0 other;

Query Match 79.9%; Score 2547; DB 21; Length 3402;  
Best Local Similarity 89.4%; Pred. No. 0;  
Matches 2930; Conservative 0; Mismatches 10; Indels 337; Gaps 4;

```

QY 89 ttcgtcgcaagtgccggtcggtgcgcacatcgcgtcagtgccagtggaagg 148
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 256 tccacgcaagtgccggtcggtgcgcacatcgcgtcagtggaagg 315
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 149 accgcgcgcgttgacatgtagacaagtgatgcgcacatccacacgcgtgagcc 208
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 316 accgcgcgcgttgacatgtagacaagtgatgcgcacatccacacgcgtgagcc 375
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 209 gcttcgcgtgctgcgcagcgaggtgaagtgaaacagtgagcgagggatgcgcgag 268
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 376 gcttcgcgtgctgcgcagcgaggtgaagtgaaacagtgagcgagggatgcgcgag 435
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 269 tgaactgtgcaagccacaaagcttcgcagcctgaagcgtaactacacccctcg 328
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 436 tgaactgtgcaagccacaaagcttcgcagcctgaagcgtaactacacccctcg 495
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 329 tgcgtgataacttagccacaggaagagagcctggggccgacagctcctctggggc 388
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 496 tgcgtgataacttagccacaggaagagagcctggggccgacagctcctctggggc 555

```

[illegible][illegible]





[illegible]

```

Db 2346 tgcgcggaacacacacgcagatgctgctcggaacacacacgcagat 2405
QY 2361 ----- 2360
Db 2406 atgctcctggacacacacagataatgctcctcaactcaacacgctgagata 2465
QY 2361 ----- 2360
Db 2466 gctggacacacacatgtgacagatatgtctggaatgacacacagtgagata 2525
QY 2361 ----- 2360
Db 2526 ctgtccggaatacacgcgcacacatgacataatgctgctggcacaactccga 2585
QY 2361 -----gcagataatgctgctggacacacacacacacagataatgctgct 2401
Db 2586 cacacatgacacacacagcagagataatgctgctggacacacacacagataatgctgct 2645
QY 2402 caaacactcacacacgctgacagatatgctcctggacacacacatgacagatatgctct 2461
Db 2646 caaacactcacacacgctgacagatatgctcctggacacacacatgacagatatgctct 2705
QY 2462 ggaacatgacacacacgctgacagatatgctcctggacacacacatgacagatat 2521
Db 2706 ggaacatgacacacacgctgacagatatgctcctggacacacacatgacagata 2765
QY 2522 tgcctgctggacacacactccgacacacatgacacacacagtgacagatatgctgct 2581
Db 2766 tgcctgctggacacacactccgacacacatgacacacacagtgacagatatgctgct 2825
QY 2582 gacacacgacagatgacgctgtgttggaagggtgctgctggaagcctgacagatgct 2641
Db 2826 gacacacgacagatgacgctgtgttggaagggtgctgctggaagcctgacagatgct 2885
QY 2642 cctggaagcctgacagatgacgctgtgttggaagggtgctgctggaagcctgacagat 2701
Db 2886 cctggaagcctgacagatgacgctgtgttggaagggtgctgctggaagcctgacagat 2945
QY 2702 ccgcctctgtccgcgcctcagctccgcgcctcagctccgcgcctcagctccgcgcct 2761
Db 2946 ccgcctctgtccgcgcctcagctccgcgcctcagctccgcgcctcagctccgcgcct 3005
QY 2762 gactatctgtccgcgcctcagctccgcgcctcagctccgcgcctcagctccgcgcct 2821
Db 3006 gactatctgtccgcgcctcagctccgcgcctcagctccgcgcctcagctccgcgcct 3065
QY 2822 ggggggacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2881
Db 3066 ggggggacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3125
QY 2882 tgcattctccctcgtacacagagagagagagagagagagagagagagagagagagag 2941
Db 3126 tgcattctccctcgtacacagagagagagagagagagagagagagagagagagagag 3185
QY 2942 atattataatgagtgagagagagagagagagagagagagagagagagagagagagag 3001
Db 3186 atattataatgagtgagagagagagagagagagagagagagagagagagagagagag 3245
QY 3002 ggaacgcgcctgtgtcttgaacgcatgtgacacacacacacacacacacacacacacac 3061
Db 3246 ggaacgcgcctgtgtcttgaacgcatgtgacacacacacacacacacacacacacacac 3305
QY 3062 cccacacacacacacacacacacacacacacacacacacacacacacacacacacacac 3121
Db 3306 cccacacacacacacacacacacacacacacacacacacacacacacacacacacacac 3365
QY 3122 agcccgatataattattattttttaaacatga 3158
Db 3366 agcccgatataattattattttttaaacacaaa 3402

```

RESULT 6  
AAFI6334

```

ID AAFI6334 strand: cDNA; 2569 BP.
XX
AC AAFI6334;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:769.
XX
KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytosolic; cardiotoxic; immunomodulatory; muscular;
KW vulnery; gastrointestinal; nephrotoxic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease; ss.
XX
OS Homo sapiens.
XX
PN WO20005174-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05988.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
PI Rosen CA, Ruben SM;
XX
DR WPI: 2000-587513/55.
DR P-PSDB: AAB57131.
XX
PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer.
XX
PS
XX
Claim 1: Page 1188-1189; 238bp; English.
XX
AAFI5566 to AAFI6505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytosolic,
CC cardiotoxic, immunomodulatory, muscular, vulnery, gastrointestinal,
CC nephrotoxic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAFI6506 to AAFI6514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
XX
SQ
Sequence 2569 BP; 588 A; 842 C; 681 G; 439 T; 19 other;
XX
Query Match 74.5%; Score 2373.6; DB 21; Length 2569;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 2535; Conservative 19; Mismatches 3; Indels 131; Gaps 4;
XX
QY 499 gtggcgaacgggacacccctcggccgacatcagtgatgaagagacgacacgcttgacg 558
Db 1 gtggcgaacgggacacccctcggccgacatcagtgatgaagagacgacacgcttgacg 60
QY 559 cgcacagagcgcgtgagccacaggaagaagaagtgacacacgacacgacacacacacacac 618
Db 61 cgcacagagcgcgtgagccacaggaagaagaagtgacacacgacacgacacacacacacac 120
QY 619 ccggaagacagcggcaaatacacctgcgcggtgtcgaacgcgcgcgcgcacatcaacgc 678
Db 121 ccggaagacagcggcaaatacacacctgcgcggtgtcgaacgcgcgcgcgcacatcaacgc 180

```



Db 2210 aagcctgagaggtgagagccatgagctagtgctacatccactgcatctccctgac 2269  
QY 2899 acagagagagggccttgatattatataagaatgaataatataatagatga 2958  
Db 2270 acagagagagggccttgatattatataagaatgaataatataatagatga 2329  
QY 2259 aggaagactggttgcaaggactgtgtctctctctgggagccgagccgtgtctt 3018  
Db 2230 aggaagactggttgcaaggactgtgtctctctctgggagccgagccgtgtctt 2389  
QY 3019 cagcagctgcatgagacacaccccgctccagagcagacacaccccccactgtcgt 3078  
Db 2290 cagcagctgcatgagacacaccccgctccagagcagacacaccccccactgtcgt 2449  
QY 3079 ggtgagcccatgcatctgttaattttagtagagttgagctgagacccgtatattat 3138  
Db 2450 ggtgagcccatgcatctgttaattttagtagagttgagctgagacccgtatattat 2509  
QY 3139 ttatttgttaacaatgaaatgcatccttccctcccaaaaaa 3186  
Db 2510 ttatttgttaacaatgaaatgcatccttccctcccaaaaaa 2557

RESULT 7

AAC76298/c  
ID AAC76298 standard; cDNA; 1737 BP.

XX AAC76298;

DT 08-FEB-2001 (first entry)

XX Human ORFX ORF1853 polynucleotide sequence SEQ ID NO:3705.

Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
vulnerary; antiproliferative; antiparkinsonian; neurotropic; neuroprotective;  
anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
hypotensive; dermatological; immunosuppressive; antineoplastic;  
antiviral; antibacterial; antifungal; antineuritic; antithyroid;  
antiangiogenic; gene therapy; cancer; proliferative disorder; hypertension;  
neurodegenerative disorder; osteoarthritis; graft vs host disease;  
cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
cholesterol ester storage; systemic lupus erythematosus; infection;  
severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
bone damage; cartilage damage; antineoplastic disease; coagulation;  
thrombosis; contraceptive; ss.

OS Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

PF 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 9905-0127607.

PR 02-APR-1999; 9905-0127636.

PR 05-APR-1999; 9905-0127728.

PR 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shimkels RA, Leach M;

XX WPI; 2000-602362/57.

DR P-PSDB; AAB42089.

XX Novel nucleic acids and peptides derived from open reading frame X,  
PT useful for treating e.g. cancers, proliferative disorders,  
XX neurodegenerative disorders and cardiovascular disease -  
PS Claim 5; Page 2856-2857; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;  
CC antiproliferative; antiparkinsonian; neurotropic; neuroprotective;  
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;  
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
CC antidiabetic; hypotensive; dermatological; immunosuppressive;  
CC antineoplastic; antibacterial; antiviral; antifungal; antineuritic;  
CC antithyroid; and antiangiogenic. The sequences can be used for determining  
CC the presence of or predisposition to, or preventing or treating  
CC pathological conditions associated with an ORFX-associated disorder. The  
CC nucleic acids can be used to express ORFX proteins in gene therapy  
CC vectors. The proteins and nucleic acids may be used to treat cancers,  
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
CC nocturnal haemoglobinuria, antineoplastic disease; to enhance  
CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 1737 BP; 318 A; 423 C; 553 G; 435 T; 8 other;

Query Match 48.0%; Score 1530.6; DB 21; Length 1737;

Best Local Similarity 98.7%; Pred. No. 0; Mismatches 4; Indels 19; Gaps 17;

Matches 1722; Conservative 0; Mismatches 4; Indels 19; Gaps 17;

QY 1446 ctgtgtgagagagctggtctccgagcag-cgccagcactactggtccagccagc 1504  
Db 1737 CTGTGTGAGGAGACATGGGTCTCCGGCAGCCGCCAGCATTTATGGCCAGCCAGT 1678  
QY 1505 tctgtgacctagtgtgtaccccaacttcaacagacatc--cacacacacacacac 1562  
Db 1677 TCCTGTGCGCTTAAGTGTACCCCAACTCTACACAGACATCCACACACACACACA 1618  
QY 1563 cacttcaacacacacttcaacagctgagagagagagagagagagagagagagag 1622  
Db 1617 CACTTTCACACACACTCACACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1558  
QY 1623 tagacgagcagctatctgcagagagagagagagagagagagagagagagagagag 1682  
Db 1557 TAGACGAGCAGCAGTATCTCAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1498  
QY 1683 atgagagagagagagctgcagagagagagagagagagagagagagagagagag 1742  
Db 1497 ATGAG 1438  
QY 1743 cccagagagagctgtgtgtgagagagagagagagagagagagagagagagagag 1802  
Db 1437 CCGAGGAGAGTCTGT 1379  
QY 1803 cctgagatgcatgtgac 1861  
Db 1378 CCGT 1319  
QY 1862 cacacagagacatgacagagatgtgcgctgtgtgtgtgtgtgtgtgtgtgtgtgt 1921  
Db 1318 CACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1259  
QY 1922 cacacagagagagacatgacagagacacacacacacacacacacacacacacacac 1979  
Db 1258 CACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1199  
QY 1980 acacccatgagagagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2038  
Db 1198 ACACCCATGAGCAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1139  
QY 2039 cgcacacagctgagagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2098  
Db 1138 C-CACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1082



```

QY 329 tgcctgagatgacatgacccagggagagagcctggggcccgacagctctctggggtc 388
    |||||||
Db 347 tgcctgagatgacatgacccagggagagagcctggggcccgacagctctctggggtc 406
QY 389 aagagagaccccgccagccagcagctgggacagacgccttcaacagccctccaagatga 448
    |||||||
Db 407 aagagagaccccgccagccagcagctgggacagacgccttcaacagccctccaagatga 466
QY 449 ggcgcgcgggtgatgcacgcgcgcctgggttagctccgctgggcttcaagtgggtggcagcg 508
    |||||||
Db 467 ggcgcgcgggtgatgcacgcgcgcctgggttagctccgctgggcttcaagtgggtggcagcg 526
QY 509 ggcacccctcgcccgccagatcagctggatgaagagacagcccgcccttaacgcgcgcagag 568
    |||||||
Db 527 ggcacccctcgcccgccagatcagctggatgaagagacagcccgcccttaacgcgcgcagag 586
QY 569 cgcctgagcccgaggaagaagaatggacactgagcctgaaacactggcgccgagagaca 628
    |||||||
Db 587 cgcctgagcccgaggaagaagaatggacactgagcctgaaacactggcgccgagagaca 646
QY 629 ggcgcaaatacacctgcccgcgtgtgcagacgcgcggcgccgacatcaacgacactacaag 688
    |||||||
Db 647 ggcgcaaatacacctgcccgcgtgtgcagacgcgcggcgccgacatcaacgacactacaag 706
QY 689 tggatgtgatccagcagcagccgcttccaaagccgctgtcacaagcagcaccgcgtgaaca 748
    |||||||
Db 707 tggatgtgatccagcagcagccgcttccaaagccgctgtcacaagcagcaccgcgtgaaca 766
QY 749 cgcacgctggagcttcggggggagacacagctctcctcaagtgcagagtgcagcagctgaagc 808
    |||||||
Db 767 cgcacgctggagcttcggggggagacacagctctcctcaagtgcagagtgcagcagctgaagc 826
QY 809 cgcctgataccagctgctggaagcgcgttgagatgagcgcgcgcgcgcgcgcgcgcgcgcgc 868
    |||||||
Db 827 cgcctgataccagctgctggaagcgcgttgagatgagcgcgcgcgcgcgcgcgcgcgcgcgc 886
QY 869 tcgaatgtggcgagcagaaattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 928
    |||||||
Db 887 tcgaatgtggcgagcagaaattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 946
QY 929 acggtctcctacacccaataagctgtctatcaaccgcgtgccgcgcgcgcgcgcgcgcgcgc 988
    |||||||
Db 947 acggtctcctacacccaataagctgtctatcaaccgcgtgccgcgcgcgcgcgcgcgcgcgc 1006
QY 989 acatctgctctggcgcaacacacatggtgtacagcttccgcgcgcgcgcgcgcgcgcgcgcgc 1048
    |||||||
Db 1007 acatctgctctggcgcaacacacatggtgtacagcttccgcgcgcgcgcgcgcgcgcgcgcgc 1066
QY 1049 tgcgcagaccccaaaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1108
    |||||||
Db 1067 tgcgcagaccccaaaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1126
QY 1109 cgtggccgcgtgtgcatcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1168
    |||||||
Db 1127 cgtggccgcgtgtgcatcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1186
QY 1169 tgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1228
    |||||||
Db 1187 tgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1246
QY 1229 ggcacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1288
    |||||||
Db 1247 ggcacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1306
QY 1289 cgcgcctcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1348
    |||||||
Db 1307 cgcgcctcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1366
QY 1349 ccacagcactactggggccagggccacagctgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1408
    |||||||
Db 1367 ccacagcactactggggccagggccacagctgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1426

```

```

QY 1409 cagacatccacacacacacac 1431
    |||||||
Db 1427 cagacatccacacacacacacac 1449

```

```

RESULT 9
AAA96724
ID AAA96724 standard; DNA; 1967 BP.
XX
XX AAA96724;
XX
XX 19-FEB-2001 (first entry)
XX
DE Polynucleotide isolated from lymph node stromal cells of fsn -/- mice.
XX
KW Lymph node stromal cell; fsn -/- mice; inflammatory disorder;
KW immune system disorder; cancer; viral infection; HIV infection;
KW blood vessel growth; tumour necrosis factor disorder; arthritis;
KW inflammatory bowel disease; fibroblast growth factor-mediated disorder;
KW cardiac failure; ss.
XX
XX Mus sp.
XX
XX Key Location/Qualifiers
XX CDS 189..1778
XX FT /*tag= a
XX
XX MO200058463-A1.
XX
XX PD 05-OCT-2000.
XX
XX PF 18-FEB-2000; 2000MO-NZ00015.
XX
XX PR 25-MAR-1999; 990US-0276268.
XX PR 26-AUG-1999; 990US-0383586.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX
XX Strachan L, Sleeman M, Abernethy N, Onrust R, Kumble KD;
XX Murison JG;
XX
XX WPI: 2000-664924/64.
XX DR P-PSDB; AAB19114.
XX
XX PT Polypeptide expressed in mammalian fsn -/- lymph node stromal cells,
XX useful for modulating growth of blood cells, for treating inflammatory
XX and tumour necrosis factor-mediated disorders, cancer and viral
XX disorders.
XX
XX PS Claim 4; Page 44-45; 75pp; English.
XX
XX The present sequence represents a polynucleotide sequence which is
XX isolated from lymph node stromal cells of fsn -/- mice. The
XX polynucleotides and their polypeptides are useful for treating an
XX inflammatory disorder, disorder of immune system and cancer selected
XX from epithelial, lymphoid, myeloid, stromal and neuronal cancers, a
XX viral disorder, in particular HIV infection and for modulating the
XX growth of blood vessels. The polypeptides are useful for treating a
XX tumour necrosis factor (TNF) mediated disorder, such as those selected
XX from arthritis, inflammatory bowel disease and cardiac failure and a
XX fibroblast growth factor-mediated disorder. It is also useful in assays
XX to determine biological activity, to raise antibodies, to isolate
XX corresponding ligands or receptors, to quantify levels of protein or
XX cognate corresponding ligand or receptors, as anti-inflammatory agents,
XX and in compositions for the treatment of skin, connective tissue and
XX immune system diseases. The polynucleotide is useful as marker for
XX tissue, as a chromosome marker or tags in the identification of a
XX genetic disorder.
XX
XX Sequence 1967 BP; 425 A; 577 C; 581 G; 384 T; 0 other;

```

Query Match 30.9%; Score 984.4; DB 21; Length 1967;











CC diagnosis, prognosis and management of colorectal cancer, breast cancer,  
CC and lung cancer. The polynucleotides can also be used to screen for  
CC peptide analogues and antagonists.

XX Sequence 551 BP; 105 A; 173 C; 136 G; 135 T; 2 other;

Query Match 17.3%; Score 550.2; DB 20; Length 551;  
Best Local Similarity 99.6%; Pred. No. 8.5e-108;  
Matches 549; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

QY 2629 gcaatcagctgctgctgagctcattatgtatgaggaattccctgctccacgctact 2688
    |||||||
DB 1 gaagtcagctgctgagctcattatgtatgaggaattccctgctccacgctact 60
QY 2689 ccccaacctgcccgcctgctccgcctcagctccgcctccacccgcctctgctc 2748
    |||||||
DB 61 ccccaacctgcccgcctgctccgcctcagctccgcctccacccgcctctgctc 120
QY 2749 cccctgctgctgctcattatgtcaccctgctgctgctcagagagctccctactgct 2808
    |||||||
DB 121 cccctgctgctgctcattatgtcaccctgctgctgctcagagagctccctactgct 180
QY 2809 gggggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 2868
    |||||||
DB 181 gggggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 240
QY 2869 ggcctacccacacacacacacacacacacacacacacacacacacacacacacacac 2928
    |||||||
DB 241 ggcctacccacacacacacacacacacacacacacacacacacacacacacacacac 300
QY 2929 agaaatgaagataataataataatgaatgaagaagactgggttgcagagactgtgct 2988
    |||||||
DB 301 agaaatgaagataataataataatgaatgaagaagactgggttgcagagactgtgct 360
QY 2989 cccctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 3048
    |||||||
DB 361 cccctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 420
QY 3049 gccagacacacacacacacacacacacacacacacacacacacacacacacacacac 3108
    |||||||
DB 421 gccagacacacacacacacacacacacacacacacacacacacacacacacacacac 480
QY 3109 gaggttgagctgaagcccgatattatatttgaataacatgaagaatgcatcctt 3168
    |||||||
DB 481 gaggttgagctgaagcccgatattatatttgaataacatgaagaatgcatcctt 540
QY 3169 tccctcccaaaa 3179
    |||||||
DB 541 tccctcccaaaa 551

```

# RESULT 13

AAA96739 ID AAA96739 standard; DNA; 1423 BP.

AC AAA96739;

DT 19-FEB-2001 (first entry)

DE Polynucleotide isolated from lymph node stromal cells of fsn -/- mice.

XX Lymph node stromal cell; fsn -/- mice; inflammatory disorder;

KW Immune system disorder; cancer; viral infection; HIV infection;

KW blood vessel growth; tumour necrosis factor disorder; arthritis;

KW inflammatory bowel disease; fibroblast growth factor-mediated disorder;

XX cardiac failure; ss.

OS Mus sp.

XX WO200058463-A1.

XX 05-OCT-2000.

XX 18-FEB-2000; 2000WO-NZ00015.

XX 25-MAR-1999; 99US-0276268.

XX 26-AUG-1999; 99US-0383586.

XX (GENE-) GENESIS RES & DEV CORP LTD.

PI Strachan L, Sleeman M, Abernethy N, Onrust R, Kumble KD;

PI Murison JG;

XX WPI: 2000-664924/64.

PT Polypeptide expressed in mammalian fsn -/- lymph node stromal cells,  
PT useful for modulating growth of blood cells, for treating inflammatory  
PT and tumour necrosis factor-mediated disorders, cancer and viral  
PT disorders

PS Claim 4; Page 64; 75pp; English.

The present sequence represents a polynucleotide sequence which is  
isolated from lymph node stromal cells of fsn -/- mice. The  
polynucleotides and their polypeptides are useful for treating an  
inflammatory disorder, disorder of immune system and cancer selected  
from epithelial, lymphoid, myeloid, stromal and neuronal cancers, a  
viral disorder, in particular HIV infection and for modulating the  
growth of blood vessels. The polypeptides are useful for treating a  
tumour necrosis factor (TNF) mediated disorder, such as those selected  
from arthritis, inflammatory bowel disease and cardiac failure and a  
fibroblast growth factor-mediated disorder. It is also useful in assays  
to determine biological activity, to raise antibodies, to isolate  
corresponding ligands or receptors, to quantify levels of protein or  
cognate corresponding ligand or receptors, as anti-inflammatory agents,  
and in compositions for the treatment of skin, connective tissue and  
immune system diseases. The polynucleotide is useful as marker for  
CC tissue, as a chromosome marker or tags in the identification of a  
CC genetic disorder.

XX Sequence 1423 BP; 306 A; 399 C; 385 G; 333 T; 0 other;

Query Match 11.7%; Score 371.4; DB 21; Length 1423;  
Best Local Similarity 77.0%; Pred. No. 9.1e-70;  
Matches 503; Conservative 0; Mismatches 71; Indels 79; Gaps 1;

```

QY 481 tccgtgagctcaaatggtggtggtggtggtggtggtggtggtggtggtggtggtggt 540
    |||||||
DB 1 tccgtgagctcaaatggtggtggtggtggtggtggtggtggtggtggtggtggtggt 60
QY 541 gacgacagagcttgacgagccagagcgctgagccacgaagaagaatgagacactg 600
    |||||||
DB 61 gatgacagagcttgacgagccagagcgctgagccacgaagaagaatgagacactg 120
QY 601 agcttgaagaactggtggtggtggtggtggtggtggtggtggtggtggtggtggt 660
    |||||||
DB 121 agcttgaagaactggtggtggtggtggtggtggtggtggtggtggtggtggtggt 180
QY 661 gctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 700
    |||||||
DB 181 gctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 240
QY 701 -----a 701
DB 241 aggacaggggcccgtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 300
QY 702 gctgacccgttccaaagccgtggtggtggtggtggtggtggtggtggtggtggtggt 761
    |||||||
DB 301 gctgacccgttccaaagccgtggtggtggtggtggtggtggtggtggtggtggtggt 360
QY 762 cgggggagacacgctccttcacgtggtggtggtggtggtggtggtggtggtggtggtggt 821
    |||||||
DB 361 cgggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 420

```

Seq	Sequence	390 BP	68 A	132 C	109 G	81 T	0 other
Query Match	11.5%;	Score 365;	DB 22;	Length 390;			
Best Local Similarity	98.7%;	Pred. No. 1.5e-68;					
Matches 368;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;			
QY	2484	atgctgtccggaatacacacgacacacatgacagatatgtcgtctggacacattcc	2543				
DB	18	acgaggtccggaatacacacgacacacatgacagatatgtcgtctggacacattcc	77				
QY	2544	ggacacacatgacacacacaggtgtgcagatatgtcgtctggacacacgacgtgacgtct	2603				
DB	78	ggacacacatgacacacacaggtgtgcagatatgtcgtctggacacacgacgtgacgtct	137				
QY	2604	tttggagaggtgtgcgtgaaacgtgcagatgctgtgcgtggaagctcatagttgataag	2663				
DB	138	tttggagaggtgtgcgtgaaacgtgcagatgctgtgcgtggaagctcatagttgataag	197				
QY	2664	ggacttccctgtctcacacgttactcccccaactctgtgcgcgtctgtccccgcctcagt	2723				
DB	198	ggacttccctgtctcacacgttactcccccaactctgtgcgcgtctgtccccgcctcaga	257				
QY	2724	cccgcctccatcccgccctctgtccctgtgcgttgcgttgcgtattttgcaactgctct	2783				
DB	258	cccgcctccatcccgccctctgtccctgtgcgttgcgttgcgtattttgcaactgctct	317				
QY	2784	gggtgtcccaagaggtccctactgctgtgtgggtgtgggggacacagcccccaagcc	2843				
DB	318	gggtgtcccaagaggtccctactgctgtgtgggtgtgggggacacagcccccaagcc	377				
QY	2844	tgaagagcttgag 2856					
DB	378	tgaagagcttgag 390					
RESULT 15							
ID	AAF27784						
XX	AAF27784 standard; cDNA; 624 BP.						
AC	AAF27784;						
XX							
DT	05-APR-2001 (first entry)						
XX							
DE	Murine MANGO 003 coding sequence SEQ ID NO: 9.						
XX							
KM	Membrane associated protein; secreted protein; human; mouse; rat;						
KM	INTERCEPT 340; MANGO 003; MANGO 347; TANGO 272; TANGO 354;						
KM	TANGO 378; skeletal disorder; cardiovascular disorder; renal disorder;						
KM	haematopoietic disorder; neural disorder; hepatic disorder;						
XX	neoplastic disease; ss.						
XX							
OS	Mus musculus.						
XX							
PN	WO200100673-A1.						
XX							
PD	04-JAN-2001.						
XX							
PF	29-JUN-2000; 2000MO-US18198.						
XX							
PR	30-JUN-1999; 99US-0345464.						
XX							
PA	(MILL-) MILLENNIUM PHARM INC.						
XX							
PI	Barnes TM, Fraser CC, Wrighton N, Myers P, Busfield SJ, Sharp JD;						
XX	WPI; 2001-050128/06.						
DR							
PT	Isolated secreted or transmembrane proteins are used for diagnosis and						
PT	treatment of neoplastic and haematopoietic disorders e.g. T cell						
PT	disorders, cancer and tumours -						
XX							
PS	Claim 1; Page 220; 294pp; English.						



ID	AAA96723	standard; DNA; 384 BP.
XX		
AC	AAA96723;	
XX		
DT	19-FEB-2001	(first entry)
XX		
DE	Polynucleotide isolated from lymph node stromal cells of fsn -/- mice.	
XX		
KW	Lymph node stromal cell; fsn -/- mice; inflammatory disorder;	
KW	Immune system disorder; cancer; viral infection; HIV infection;	
KW	blood vessel growth; tumour necrosis factor disorder; arthritis;	
KW	inflammatory bowel disease; fibroblast growth factor-mediated disorder	
KX	cardiac failure; ss.	
XX		
OS	Mus sp.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	149..992
FT		/*tag="a
FT		/transl_except= (pos: 368..373, aa: Xaa)
FT		/note= "Xaa is an unspecified amino acid"
XX		
PN	WO200058463-A1.	
XX		
PD	05-OCT-2000.	
XX		
PF	18-FEB-2000; 2000WO-NZ00015.	
XX		
PR	25-MAR-1999; 99US-0276268.	
PR	26-AUG-1999; 99US-0383586.	
XX		
PA	(GENE-) GENESIS RES & DEV CORP LTD.	
XX		
PI	Strachan L, Sleeman M, Abernethy N, Onrust R, Kumble KD;	
PI	Murison JG;	
XX		
DR	WPI: 2000-664924/64.	
DR	P-PSDB: AMB19113.	
XX		
FT	Polypeptide expressed in mammalian fsn -/- lymph node stromal cells,	
FT	useful for modulating growth of blood cells, for treating inflammatory	
PT	and tumour necrosis factor-mediated disorders, cancer and viral	
PT	disorders	
XX		
PS	Claim 4; Page 44; 75pp; English.	
XX		
CC	The present sequence represents a polynucleotide sequence which is	
CC	isolated from lymph node stromal cells of fsn -/- mice. The	
CC	polynucleotides and their polypeptides are useful for treating an	
CC	inflammatory disorder, disorder of immune system and cancer selected	
CC	from epithelial, lymphoid, myeloid, stromal and neuronal cancers, a	
CC	viral disorder, in particular HIV infection and for modulating the	
CC	growth of blood vessels. The polypeptides are useful for treating a	
CC	tumour necrosis factor (TNF) mediated disorder, such as those selected	
CC	from arthritis, inflammatory bowel disease and cardiac failure and a	
CC	fibroblast growth factor-mediated disorder. It is also useful in assays	
CC	to determine biological activity, to raise antibodies, to isolate	
CC	corresponding ligands or receptors, to quantify levels of protein or	
CC	cognate corresponding ligand or receptors, as antiinflammatory agents,	
CC	and in compositions for the treatment of skin, connective tissue and	
CC	immune system diseases. The polynucleotide is useful as marker for	
CC	tissue, as a chromosome marker or tags in the identification of a	
CC	genetic disorder.	
XX		
Q0	Sequence 384 BP; 72 A; 113 C; 111 G; 87 T; 1 other;	

	Query Match	9.8%;	Score 312.6;	DB 21;	Length 384;
	Best Local Similarity	88.3%;	Pred. NO. 1.9e-57;		
	Matches 339;	Conservative 0;	Mismatches 45;	Indels 0;	Gaps 0;
QY 753	ggtgacttcgggggagcaacgctccctcccaagtgcagaagtgcacagcgtcaagccggt				

[illegible]

RESULT 18	
AA214137	
ID	AA214137 standard; cDNA; 300 BP.
XX	
AC	AA214137;
XX	
DT	12-OCT-1999 (first entry)
XX	
DE	Human gene expression product cDNA sequence SEQ ID NO:1606.
XX	
HW	Human; gene; gene expression product; diagnosis; therapy; probe;
KW	detection; mapping; tissue typing; profiling; forensic; cancer;
KW	genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX	
OS	Homo sapiens.
XX	
PN	MO9938972-A2.
XX	
PD	05-AUG-1999.
XX	
PE	28-JAN-1999; 99WO-US01619.
XX	
PR	03-APR-1998; 98US-0080665.
PR	28-JAN-1998; 98US-0072910.
PR	24-FEB-1998; 98US-0075954.
PR	31-MAR-1998; 98US-0080114.
PR	03-APR-1998; 98US-0080515.
XX	
PA	(CHIR ) CHIRON CORP.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
PI	Escobedo J, Garcia PD, Garcia V, Glese K, Innis MA;
PI	Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
PI	Lamson G, Lesnikowitz D, Pot D, Randazzo F, Reinhard C;
PI	Stache-Crain B, Suduth-Klinger J, Williams LT;
XX	
DR	WPI; 1999-494092/41.
XX	
PT	Novel human genes and their expression products which are
PT	differentially expressed in different cell types
XX	
PS	Claim 1; Page 1024; 2479pp; English.
XX	
CC	The present invention describes a library of human polynucleotides
CC	comprising the sequences given in AA212532 to AA217779. Also described is











OY 832 gtaggg---gtacggggccgagggccgacacatccacatcgtgagtgggcgcagaag 888  
 ||||| | ||||| | |||| | |||| | |||| | |||| | ||||  
 Db 871 gggggggtagaacgcagcagaaggtgggcccggagcagcagcaccctacgtgcctcaag 930  
 OY 889 t 889  
 |  
 Db 931 t 931  
  
 RESULT 25  
 ID AAAS4413  
 AC AAAS4413 standard; DNA; 2427 BP.  
 XX AAAS4413;  
 DT 11-APR-2001 (first entry)  
 DE Mutant (K652M) fibroblast growth factor receptor 3 (FGFR3-IIIB).  
 XX Fibroblast growth factor 3 receptor; FGFR3; mutant; detection;  
 KW cancer; carcinoma; lung cancer; breast cancer; colon cancer;  
 KW skin cancer; bladder; cervix; human; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200068424-A2.  
 XX  
 PD 16-NOV-2000.  
 XX  
 PF 04-MAY-2000; 2000WO-EP04591.  
 XX  
 PR 05-MAY-1999; 99US-0132705.  
 XX  
 PA (CURT-) INST CURIE.  
 XX (CNRS ) CNRS CENT NAT RECH SCL.  
 XX  
 PI Cappellen D, Chopin D, Radvanyi F, Ricol D, Thery J;  
 XX  
 XX WPI; 2001-016103/02.  
 DR  
 XX  
 PT Diagnosing carcinoma e.g. bladder or cervix carcinomas in a biological  
 PT sample such as tissue, bone marrow or body fluid, preferably from 3  
 PT animal or human, by identifying fibroblast growth factor receptor 3  
 PT mutations  
 XX  
 PS  
 PS Disclosure; Fig 2h; 41pp; English.  
 XX  
 CC The identification of fibroblast growth factor receptor 3 (FGFR3)  
 CC mutations in a biological sample such as tissue, bone marrow or  
 CC body fluid e.g. urine, from a warm-blooded animal, preferably human  
 CC is useful for diagnosing carcinomas such as human bladder and cervix  
 CC carcinomas, or cancers associated with lung, breast, colon and skin.  
 CC The pharmaceutical preparations comprising agents which inhibit the  
 CC synthesis and expression of FGFR3 and so have an anti-proliferation  
 CC effect on carcinomas can be used to treat cancer.  
 XX  
 XX  
 Sequence 2427 BP; 427 A; 784 C; 821 G; 395 T; 0 other;

[illegible][illegible]

RESULT 26

ID AAA54414 standard; DNA: 2427 BP.

AC AAA54414;

DT 11-APR-2001 (first entry)

DE Mutant (X809C) fibroblast growth factor receptor 3 (FGFR3-IIIB).

KW Fibroblast growth factor 3 receptor; FGFR3; mutant; detection;  
cancer; carcinoma; lung cancer; breast cancer; colon cancer;  
skin cancer; bladder; cervix; human; ds.

OS Homo sapiens.

PN W0200068424-A2.

PD 16-NOV-2000.

PF 04-MAY-2000; 2000MO-EP04591.

PR 05-MAY-1999; 99US-0132705.

PA (CURT-) INST CURIE.  
(CNRS ) CNRS CENT NAT RECH SCI.

PI Cappellen D, Chopin D, Radvanyi F, Ricol D, Thlery J;  
WPT: 2001-016103/02.

PT Diagnosing carcinoma e.g. bladder or cervix carcinomas in a biological  
sample such as tissue, bone marrow or body fluid, preferably from  
animal or human, by identifying fibroblast growth factor receptor 3  
mutations

PS Disclosure; Fig 2i; 41pp; English.

XX The identification of fibroblast growth factor receptor 3 (FGFR3)  
mutations in a biological sample such as tissue, bone marrow or  
body fluid e.g. urine, from a warm-blooded animal, preferably human  
is useful for diagnosing carcinomas such as human bladder and cervix  
carcinomas, or cancers associated with lung, breast, colon and skin.  
The pharmaceutical preparations comprising agents which inhibit the  
synthesis and expression of FGFR3 and so have an anti-proliferation

CC effect on carcinomas can be used to treat cancer.  
XX  
SQ Sequence 2427 BP; 427 A; 785 C; 821 G; 394 T; 0 other;

Query Match 3.7%; Score 118.6; DB 22; Length 2427;  
Best Local Similarity 56.1%; Pred. No. 4.3e-16;  
Matches 270; Conservative 0; Mismatches 199; Indels 12; Gaps 2;

```
OY 421 ccgcgttcaacagccctccaagatgagcgccggtgtatctgcacgcccgtggtagc 480
    || || || || || || || || || || || || || || || || || || || || ||
DB 451 ccttactgagcacgcccgcgagtgagacaagaagctgtgcctgcgtgcgcgcgccaac 510
OY 481 tccgtgcgtcgaagtgtgtgcccagcgccctccgcgcgacatacgtgtgataag 540
    || || || || || || || || || || || || || || || || || || || || ||
DB 511 accgttcgcttcgctgtccagcgtctgcaaccacatccctccatctcctgtgtgaag 570
OY 541 gacgacacagcctt-----gacgcgcacagagcgcgctgagcccaaggaagaag 591
    || || || || || || || || || || || || || || || || || || || || ||
DB 571 aacggcagggagttccgcgcgagcaaccgcatgtgaggtcaagtctgcgcatcaagcag 630
OY 592 tggacacatgagccttgaagaacctgcgcgagagacagcgcaatacactgcgcgtg 651
    || || || || || || || || || || || || || || || || || || || || ||
DB 631 tggagcctgtgtcatgtgaaagcgtgtgtccctcgcgcgcgcaactacactgtcgtgtg 690
OY 652 tcgaaccgcgcggcgccatcaacgccaactacaagttgagtgtatcagcggaccggt 711
    || || || || || || || || || || || || || || || || || || || || ||
DB 691 gagaacaagtttgcgcagcatccgcgcagactacacgttgaacgtgtcgtgagcgctcccg 750
OY 712 tccaagccgctgtctcaacagcagcccccgtgtgaacacgagcggttgacttcgggggacc 771
    || || || || || || || || || || || || || || || || || || || || ||
DB 751 caccgcgccatcctcgacagcgggcgctgcgcgccaacacagcgcggtgcgtgcgacgagac 810
OY 772 acgttccttcagtgcaaggtgcgcagcagcgttgaagccggtgtatccagttgcgaaagcgc 831
    || || || || || || || || || || || || || || || || || || || || ||
DB 811 gtgagttcactcgaaggtgtacagtgacgacagccccaactcaactcagttgctcaagcac 870
OY 832 gtgtga---gtacggcgccgagggcgcccaacaactccaactcgtatgtggcgccagaag 888
    || || || || || || || || || || || || || || || || || || || || ||
DB 871 gtgaggtgagcagcgacgaggtgtggcccgacgacgacacacacttaccgtgtctaaag 930
OY 889 t 889
DB 931 t 931
```

## RESULT 27

AAA54415  
ID AAA54415 standard; DNA; 2427 BP.

AC AAA54415;

DT 11-APR-2001 (first entry)

DE Mutant (X809G) fibroblast growth factor receptor 3 (FGFR3-IIID).

KW Fibroblast growth factor 3 receptor; FGFR3; mutant; detection;  
KW cancer; carcinoma; lung cancer; breast cancer; colon cancer;  
KW skin cancer; bladder; cervix; human; ds.

OS Homo sapiens.

PN WO200068424-A2.

PD 16-NOV-2000.

PF 04-MAY-2000; 2000WO-EP04591.

PR 05-MAY-1999; 99US-0132705.

PA (CURI-) INST CURIE.

PA (CNRS ) CNRS CENT NAT RECH SCI.

XX

PI Cappellen D, Chopin D, Radvanyi F, Rioul D, Thierly J;  
XX  
DR WPI; 2001-016103/02.  
XX

PT Diagnosing carcinoma e.g. bladder or cervix carcinomas in a biological  
sample such as tissue, bone marrow or body fluid, preferably from  
PT animal or human, by identifying fibroblast growth factor receptor 3  
PT mutations

PS Disclosure; Fig 2j; 41pp; English.

XX  
XX The identification of fibroblast growth factor receptor 3 (FGFR3)  
CC mutations in a biological sample such as tissue, bone marrow or  
CC body fluid e.g. urine, from a warm-blooded animal, preferably human  
CC is useful for diagnosing carcinomas such as human bladder and cervix  
CC carcinomas, or cancers associated with lung, breast, colon and skin.  
CC The pharmaceutical preparations comprising agents which inhibit the  
CC synthesis and expression of FGFR3 and so have an anti-proliferation  
CC effect on carcinomas can be used to treat cancer.

XX  
SQ Sequence 2427 BP; 428 A; 784 C; 822 G; 393 T; 0 other;

Query Match 3.7%; Score 118.6; DB 22; Length 2427;  
Best Local Similarity 56.1%; Pred. No. 4.3e-16;  
Matches 270; Conservative 0; Mismatches 199; Indels 12; Gaps 2;

```
OY 421 ccgcgttcaacagccctccaagatgagcgccggtgtatctgcacgcccgtggtagc 480
    || || || || || || || || || || || || || || || || || || || || ||
DB 451 ccttactgagcacgcccgcgagtgagacaagaagctgtgcctgcgtgcgcgcgccaac 510
OY 481 tccgtgcgtcgaagtgtgtgcccagcgccctccgcgcgacatacgtgtgataag 540
    || || || || || || || || || || || || || || || || || || || || ||
DB 511 accgttcgcttcgctgtccagcgtctgcaaccacatccctccatctcctgtgtgaag 570
OY 541 gacgacacagcctt-----gacgcgcacagagcgcgctgagcccaaggaagaag 591
    || || || || || || || || || || || || || || || || || || || || ||
DB 571 aacggcagggagttccgcgcgagcaaccgcatgtgaggtcaagtctgcgcatcaagcag 630
OY 592 tggacacatgagccttgaagaacctgcgcgagagacagcgcaatacactgcgcgtg 651
    || || || || || || || || || || || || || || || || || || || || ||
DB 631 tggagcctgtgtcatgtgaaagcgtgtgtccctcgcgcgcgcaactacactgtcgtgtg 690
OY 652 tcgaaccgcgcggcgccatcaacgccaactacaagttgagtgtatcagcggaccggt 711
    || || || || || || || || || || || || || || || || || || || || ||
DB 691 gagaacaagtttgcgcagcatccgcgcagactacacgttgaacgtgtcgtgagcgctcccg 750
OY 712 tccaagccgctgtctcaacagcagcccccgtgtgaacacgagcggttgacttcgggggacc 771
    || || || || || || || || || || || || || || || || || || || || ||
DB 751 caccgcgccatcctcgacagcgggcgctgcgcgccaacacagcgcggtgcgtgcgacgagac 810
OY 772 acgttccttcagtgcaaggtgcgcagcagcgttgaagccggtgtatccagttgcgaaagcgc 831
    || || || || || || || || || || || || || || || || || || || || ||
DB 811 gtgagttcactcgaaggtgtacagtgacgacagccccaactcaactcagttgctcaagcac 870
OY 832 gtgtga---gtacggcgccgagggcgcccaacaactccaactcgtatgtggcgccagaag 888
    || || || || || || || || || || || || || || || || || || || || ||
DB 871 gtgaggtgagcagcgacgaggtgtggcccgacgacgacacacacttaccgtgtctaaag 930
OY 889 t 889
DB 931 t 931
```

## RESULT 28

AAA54416  
ID AAA54416 standard; DNA; 2427 BP.

AC AAA54416;

DT 11-APR-2001 (first entry)

DE Mutant (X809G) fibroblast growth factor receptor 3 (FGFR3-IIID).

XX Fibroblast growth factor 3 receptor; FGFR3; mutant; detection;  
 KW cancer; carcinoma; lung cancer; breast cancer; colon cancer;  
 KM skin cancer; bladder; cervix; human; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO20006424-A2.  
 XX  
 PD 16-NOV-2000.  
 XX  
 PF 04-MAY-2000; 2000WO-EP04591.  
 XX  
 PR 05-MAY-1999; 99US-0132705.  
 XX  
 PA (CURI-) INST CURIE.  
 XX  
 PI (CNRS ) CNRS CENT NAT RECH SCI.  
 XX  
 PI Cappellen D, Chopin D, Radvanyi F, Ricol D, Thierly J;  
 XX  
 DR WPI; 2001-016103/02.  
 XX  
 PT Diagnosing carcinoma e.g. bladder or cervix carcinomas in a biological  
 PT sample such as tissue, bone marrow or body fluid, preferably from  
 PT animal or human, by identifying fibroblast growth factor receptor 3  
 PT mutations  
 XX  
 PS Disclosure; Fig 2K; 41pp; English.  
 XX  
 CC The identification of fibroblast growth factor receptor 3 (FGFR3)  
 CC mutations in a biological sample such as tissue, bone marrow or  
 CC body fluid e.g. urine, from a warm-blooded animal, preferably human  
 CC is useful for diagnosing carcinomas such as human bladder and cervix  
 CC carcinomas, or cancers associated with lung, breast, colon and skin.  
 CC The pharmaceutical preparations comprising agents which inhibit the  
 CC synthesis and expression of FGFR3 and so have an anti-proliferation  
 CC effect on carcinomas can be used to treat cancer.  
 CC  
 SQ Sequence 2427 BP; 429 A; 784 C; 821 G; 393 T; 0 other;

Query Match 3.7%; Score 118.6; DB 22; Length 2427;  
 Best Local Similarity 56.1%; Pred. No. 4.3e-16;  
 Matches 270; Conservative 0; Mismatches 199; Indels 12; Gaps 2;  
 QY 421 ccgcgcctcacacagccctccaagatgagcgcggtgatcgacgagccgtggtacg 480  
 DB 451 ccttactgacacagcgccgagcgatgacagaagctgtgctgcccgtgcccgcacac 510  
 QY 481 tccgtgagcgtcaagtgtgctgacgagcgacccctcgccgacatacagtgatgaag 540  
 DB 511 accgttcgccttcgctgcgcacgctgacacccacccatccctcatctcgtgctgaag 570  
 QY 541 gacgcacagcctt-----gacgcgcccagagcgcgctgagccaggaagaag 591  
 DB 571 aacggcagggaatttcgcggcgagcagccgcatgtggaagatcaagctcggcataagca 630  
 QY 592 tggacactgagctgaagaacctcgccggcgaggaacagcggaataacatctgcgcgtg 651  
 DB 631 tggagcctgtgcatggaagcgtgtgtccctcgagcgcggaactacacctcgtcgtg 690  
 QY 652 tggaacccgagcggtgcataacgcacacccatacaggtgtgagtgtatccagcgaccgt 711  
 DB 691 gagaacaagtttggcagatccgcgacagatcaacgctgacgtgtgtgagcgctcccg 750  
 QY 712 tccaaagccgtgtcagcagcagcagcccgtaacacagcgtgtgacttcggggagac 771  
 DB 751 caccggcccatctctgcagcgaggtgtgcccgaacacgaagcggtgtctggtgacgac 810  
 QY 772 acgtcttcacagtgcaaggtgcgacgacgtgaagccggtgatccagtggtgaagcg 831  
 DB 811 gtgagttcactcactgaaggtgtacagtgaacagccccaatccagtggtcctaagac 870

QY 832 gtgga---gtacggcgccgagggcgcccaactccacatgatgtggcgccagaag 888  
 DB 871 gtggaggtgaagcgacgaaggtggccgagcagcacaacccaatgatacgtgtcgaag 930  
 QY 889 t 889  
 DB 931 t 931

## RESULT 29

AAAS4417  
 ID AAAS4417 standard; DNA; 2427 BP.  
 XX  
 AC AAAS4417;  
 XX  
 DT 11-APR-2001 (first entry)  
 XX  
 DE Mutant (X809G) fibroblast growth factor receptor 3 (FGFR3-IIIB).  
 XX  
 KW Fibroblast growth factor 3 receptor; FGFR3; mutant; detection;  
 KW cancer; carcinoma; lung cancer; breast cancer; colon cancer;  
 KW skin cancer; bladder; cervix; human; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO20006424-A2.  
 XX  
 PD 16-NOV-2000.  
 XX  
 PF 04-MAY-2000; 2000WO-EP04591.  
 XX  
 PR 05-MAY-1999; 99US-0132705.  
 XX  
 PA (CURI-) INST CURIE.  
 XX  
 PI (CNRS ) CNRS CENT NAT RECH SCI.  
 XX  
 PI Cappellen D, Chopin D, Radvanyi F, Ricol D, Thierly J;  
 XX  
 DR WPI; 2001-016103/02.  
 XX  
 PT Diagnosing carcinoma e.g. bladder or cervix carcinomas in a biological  
 PT sample such as tissue, bone marrow or body fluid, preferably from  
 PT animal or human, by identifying fibroblast growth factor receptor 3  
 PT mutations  
 XX  
 PS Disclosure; Fig 2L; 41pp; English.  
 XX  
 CC The identification of fibroblast growth factor receptor 3 (FGFR3)  
 CC mutations in a biological sample such as tissue, bone marrow or  
 CC body fluid e.g. urine, from a warm-blooded animal, preferably human  
 CC is useful for diagnosing carcinomas such as human bladder and cervix  
 CC carcinomas, or cancers associated with lung, breast, colon and skin.  
 CC The pharmaceutical preparations comprising agents which inhibit the  
 CC synthesis and expression of FGFR3 and so have an anti-proliferation  
 CC effect on carcinomas can be used to treat cancer.  
 CC  
 SQ Sequence 2427 BP; 428 A; 785 C; 821 G; 393 T; 0 other;

Query Match 3.7%; Score 118.6; DB 22; Length 2427;  
 Best Local Similarity 56.1%; Pred. No. 4.3e-16;  
 Matches 270; Conservative 0; Mismatches 199; Indels 12; Gaps 2;  
 QY 421 ccgcgcctcacacagccctccaagatgagcgcggtgatcgacgagccgtggtacg 480  
 DB 451 ccttactggaacagcgccgagcgatggaagaagctgtgctgcccgtgcccgcacac 510  
 QY 481 tccgtgagcgtcaagtgtgctgacgagcgacccctcgccgacatacagtgatgaag 540  
 DB 511 accgttcgccttcgctgcgcacgctgacacccacccatccctcatctcgtgctgaag 570  
 QY 541 gacgcacagcctt-----gacgcgcccagagcgcgctgagccaggaagaag 591

```
Db 571 aacggcagggaggttcgcgagcaccgcatttgaggacatcaagctgcgcacacag 630
Qy 592 tggacactgagcctgaaagaacctgcgcgagagacagcggcaaatcaacctgcgcgtg 651
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 631 tggagcctgtgtcatggaagcgtgtgtccctcgaccgcgcgaactacaacctgcgtcgtg 690
Qy 652 tcgagacgcgcggcgccatcaacgacacgaagtgatgtgatccgcgcgcgcgt 711
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 691 gagacaagtttgacagcatccgcagacgtaacagctgcgtgcgtgacgctcccg 750
Qy 712 tccaagccgtgcgcacagacgaccccgctgacacagcagtgacgttcggggggacc 771
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 751 caccggcccatcctgcagcggcggtgtgcgcgcacacagacggcggtgcgtgcagcgac 810
Qy 772 acgtccttcagtgcaaggtgcgcagcagctgaagccggtgatccagtgctgaagcgc 831
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 811 gtggagttccactgcaaggtgtacagtgcacgacaccccatccagtgctcaagc 870
Qy 832 gtgga---gtacggcgccgcgagggcgccacacatccatccatgtgtgggggcagag 888
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 871 gtggaggttgaaagcgaaggtgtggcccgagcgcgcacacccatgtaaccgtgtcaag 930
Qy 889 t 889
    |
Db 931 t 931

RESULT 30
ID AAA54418 standard; DNA: 2427 BP.
XX AAA54418;
AC
XX
XX 11-APR-2001 (first entry)
DE Mutant (X809L) fibroblast growth factor receptor 3 (FGFR3-IIIB).
XX Fibroblast growth factor 3 receptor; FGFR3; mutant; detection;
KW cancer; carcinoma; lung cancer; breast cancer; colon cancer;
KM skin cancer; bladder; cervix; human; ds.
XX Homo sapiens.
OS
XX
XX WO200068424-A2.
PN
XX 16-NOV-2000.
PD
XX
XX 04-MAY-2000; 2000WO-EP04591.
PE
XX 05-MAY-1999; 99US-0132705.
PR
XX (CURI-) INST CURIE.
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
XX Cappellet D, Chopin D, Radvanyi F, Riccol D, Thierry J;
PI WPI; 2001-016103/02.
DR
XX
XX Diagnosing carcinoma e.g. bladder or cervix carcinomas in a biological
PT sample such as tissue, bone marrow or body fluid, preferably from
PT animal or human, by identifying fibroblast growth factor receptor 3
PT mutations
XX
XX Disclosure; Fig 2m; 41pp; English.
XX
XX The identification of fibroblast growth factor receptor 3 (FGFR3)
XX mutations in a biological sample such as tissue, bone marrow or
XX body fluid e.g. urine, from a warm-blooded animal, preferably human
XX is useful for diagnosing carcinomas such as human bladder and cervix
XX carcinomas, or cancers associated with lung, breast, colon and skin.
XX The pharmaceutical preparations comprising agents which inhibit the
XX synthesis and expression of FGFR3 and so have an anti-proliferation
XX effect on carcinomas can be used to treat cancer.
XX
```

```
XX
SQ Sequence 2427 BP; 428 A; 784 C; 820 G; 395 T; 0 other;
Query Match 3.7%; Score 118.6; DB 22; Length 2427;
Best Local Similarity 56.1%; Pred. No. 4.3e-16;
Matches 270; Conservative 0; Mismatches 199; Indels 12; Gaps 2;
Qy 421 ccgcgcttcacacagccctccaagatgagcggcggtgatcgacgagcccggtgtagc 480
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 451 cctcatggaacagccgcgagcagatgtgacaagaagctgtgcgtgcgcgcgcgcgcac 510
Qy 481 tcggtggcctcaaggtgcgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 511 accgtccgcttcgcgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 570
Qy 541 gacacacagcctc-----gacgcgcacagagccgcgtgacccacgaagaagaag 591
    || | || | || | || | || | || | || | || | || | || | || | || | || |
Db 571 aacggcagggaggttcgcgagcagcaccgcatltggaggatcaagctgcgcgcacagcag 630
Qy 592 tggacactgagcctgaaagaacctgcgcgagagacagcggcaaatcaacctgcgcgtg 651
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 631 tggagcctgtcatggaagcgtgtgtgtccctcgaccgcgcgaactacaacctgcgtcgtg 690
Qy 652 tcgaaacgcgcggcgccatcaacgacacatcaaggtgtgtgtatccgcgcgcgcgt 711
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 691 gagacaagtttgacagcatccgcagacgtaacagctgcgtgcgtgagcgtcccg 750
Qy 712 tccaagccgtgcgcacagacgaccccgctgacacagcagtgacgttcggggggacc 771
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 751 caccggcccatcctgcagcggcggtgtgcgcgcacacagacggcggtgcgtgcagcgac 810
Qy 772 acgtccttcagtgcaaggtgcgcagcagctgaagccggtgatccagtgctgaagcgc 831
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 811 gtggagttccactgcaaggtgtacagtgcacgacaccccatccagtgctcaagc 870
Qy 832 gtgga---gtacggcgccgcgagggcgccacacatccatccatgtgtgggggcagag 888
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 871 gtggaggttgaaagcgaaggtgtggcccgagcgcgcacacccatgtaaccgtgtcaag 930
Qy 889 t 889
    |
Db 931 t 931

RESULT 31
ID AAA54419 standard; DNA: 2427 BP.
XX AAA54419
AC
XX
XX 11-APR-2001 (first entry)
DE Mutant (N542K) fibroblast growth factor receptor 3 (FGFR3-IIIB).
XX Fibroblast growth factor 3 receptor; FGFR3; mutant; detection;
KW cancer; carcinoma; lung cancer; breast cancer; colon cancer;
KM skin cancer; bladder; cervix; human; ds.
XX Homo sapiens.
OS
XX
XX WO200068424-A2.
PN
XX 16-NOV-2000.
PD
XX
XX 04-MAY-2000; 2000WO-EP04591.
PE
XX 05-MAY-1999; 99US-0132705.
PR
XX (CURI-) INST CURIE.
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
XX Cappellet D, Chopin D, Radvanyi F, Riccol D, Thierry J;
PI
```

XX WPI; 2001-016103/02.  
DR  
XX  
XX  
PT Diagnosing carcinoma e.g. bladder or cervix carcinomas in a biological  
PT sample such as tissue, bone marrow or body fluid, preferably from  
PT animal or human, by identifying fibroblast growth factor receptor 3  
PT mutations  
XX  
PS  
PS Disclosure; Fig 2n; 41pp; English.  
XX  
XX The identification of fibroblast growth factor receptor 3 (FGFR3)  
CC mutations in a biological sample such as tissue, bone marrow or  
CC body fluid e.g. urine, from a warm-blooded animal, preferably human  
CC is useful for diagnosing carcinomas such as human bladder and cervix  
CC carcinomas, or cancers associated with lung, breast, colon and skin.  
CC The pharmaceutical preparations comprising agents which inhibit the  
CC synthesis and expression of FGFR3 and so have an anti-proliferation  
CC effect on carcinomas can be used to treat cancer.  
XX  
XX Sequence 2427 BP; 429 A; 783 C; 821 G; 394 T; 0 other;

Query Match	3.7%	Score 118.6;	DB 22;	Length 2427;
Best Local Similarity	56.1%	Pred. No. 4.3e-16;		
Matches 270; Conservative	0;	Mismatches 199;	Indels 12;	Gaps 2

KW	Fibroblast growth factor 3 receptor; EGFR3; mutant; detection;
KV	cancer; carcinoma; lung cancer; breast cancer; colon cancer;
KM	skin cancer; bladder; cervix; human; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200068424-A2.
XX	
PD	16-NOV-2000.
XX	
PF	04-MAY-2000; 2000WO-EP04591.
XX	
PR	05-MAY-1999; 99US-0132705.
XX	
PA	(CURT-) INST CURIE.
PA	(CNRS ) CNRS CENT NAT RECH SCI.
XX	
PI	Cappellen D, Chopin D, Radvanyi F, Ricol D, Thery J;
DR	WPI: 2001-016103/02.
XX	
PT	Diagnosing carcinoma e.g. bladder or cervix carcinomas in a biological
PT	sample such as tissue, bone marrow or body fluid, preferably from
PT	animal or human, by identifying fibroblast growth factor receptor 3
PT	mutations

[illegible]

PS Disclosure; Fig 20; 41pp; English.

XX

CC The identification of fibroblast growth factor receptor 3 (FGFR3)

CC mutations in a biological sample such as tissue, bone marrow or

CC body fluid e.g. urine, from a warm-blooded animal, preferably human

CC is useful for diagnosing carcinomas such as human bladder and cervix

CC carcinomas, or cancers associated with lung, breast, colon and skin.

CC The pharmaceutical preparations comprising agents which inhibit the

CC synthesis and expression of FGFR3 and so have an anti-proliferation

CC effect on carcinomas, can be used to treat cancer.

XX

50 . Sequence 2427 BP; 428 A; 783 C; 822 G; 394 T; 0 other;

[illegible][illegible]

RESULT	32
AAA54420	
ID	AAA54420 standard; DNA; 2427 BP.
XX	
AC	AAA54420;
XX	
DT	11-APR-2001 (first entry)
XX	
DE	Mutant (N542K) fibroblast growth factor receptor 3 (FCGR3-IIb),

[illegible]

```
Db 871 gtggagggtgaacgagcagaaggtggcccgagcgacacacctgcttaccgtgtctcaag 930
QY 889 t 889
Db 931 t 931

RESULT 33
AAA54421
ID AAA54421 standard; DNA: 2427 BP.
XX
XX AAA54421:
AC
XX 11-APR-2001 (first entry)
XX
DE Mutant (G382R) fibroblast growth factor receptor 3 (FGFR3-IIId).
XX
KW Fibroblast growth factor 3 receptor; FGFR3; mutant; detection;
KW cancer; carcinoma; lung cancer; breast cancer; colon cancer;
KW skin cancer; bladder; cervix; human; ds.
XX
OS Homo sapiens.
XX
PN WO200068424-A2.
XX
PD 16-NOV-2000.
XX
XX 04-MAY-2000; 2000WO-EP04591.
XX
XX 05-MAY-1999; 99US-0132705.
XX
PA (CURT-) INST CURIE.
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI Cappellen D, Chopin D, Radvanyi F, Ricol D, Thierry J;
XX
XX WPI; 2001-016103/02.
XX
DR
XX
XX Diagnosing carcinoma e.g. bladder or cervix carcinomas in a biological
XX sample such as tissue, bone marrow or body fluid, preferably from
XX animal or human, by identifying fibroblast growth factor receptor 3
XX mutations
XX
PS Disclosure; Fig 2p; 41pp; English.
XX
XX The identification of fibroblast growth factor receptor 3 (FGFR3)
XX mutations in a biological sample such as tissue, bone marrow or
XX body fluid e.g. urine, from a warm-blooded animal, preferably human
XX is useful for diagnosing carcinomas such as human bladder and cervix
XX carcinomas, or cancers associated with lung, breast, colon and skin.
XX The pharmaceutical preparations comprising agents which inhibit the
XX synthesis and expression of FGFR3 and so have an anti-proliferation
XX effect on carcinomas can be used to treat cancer.
XX
XX Sequence 2427 BP; 429 A; 784 C; 820 G; 394 T; 0 other;

Query Match 3.7%; Score 118.6; DB 22; Length 2427;
Best Local Similarity 56.1%; Pred. No. 4.3e-16;
Matches 270; Conservative 0; Mismatches 199; Indels 12; Gaps 2;
```

```
QY 592 tggacactgagctcgtgaagaacctgcgcccggaggacagcggaataacacctgcgcgtg 651
Db 631 tggagcctgtgtaagaaagcgtgtgcccctcgagccgcggaactaacaacctgcgtcgtg 690
QY 652 tcgacccgcgcggcgccatcaacgcccactacaagtgtagtgcacagcgagccgt 711
Db 691 gagaacaagtttgcagcatccgcagatgcagatgcacgcgtcgcgtgcgagcgtcccg 750
QY 712 tccaagccctgtctacaagcagcagcccgctgaacaacagcgttgacctcgggggacc 771
Db 751 caccgcccacatccctgcagcgcggcgtgcgcgcaaccagacgcgtgtcgtgcagcgac 810
QY 772 acgctcttcagtgcaaggttcgcagcagctgtagccggtgataccagtgctgaagcgc 831
Db 811 gtggagttccactgcaaggtgtacagtcagcaagcccacatcagttggtctaaagac 870
QY 832 gtgga---gtacgcgcgcgaagcgccacaactccacatcagttgtggcgccagaag 888
Db 871 gtggaggtgaacgagcagaaggtggcccgagcgacacacctagttaccgtgtctcaag 930
QY 889 t 889
Db 931 t 931

RESULT 34
AAA54422
ID AAA54422 standard; DNA: 2427 BP.
XX
XX AAA54422:
AC
XX 11-APR-2001 (first entry)
XX
DE Mutant (G382R) fibroblast growth factor receptor 3 (FGFR3-IIId).
XX
KW Fibroblast growth factor 3 receptor; FGFR3; mutant; detection;
KW cancer; carcinoma; lung cancer; breast cancer; colon cancer;
KW skin cancer; bladder; cervix; human; ds.
XX
OS Homo sapiens.
XX
PN WO200068424-A2.
XX
PD 16-NOV-2000.
XX
XX 04-MAY-2000; 2000WO-EP04591.
XX
XX 05-MAY-1999; 99US-0132705.
XX
PA (CURT-) INST CURIE.
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI Cappellen D, Chopin D, Radvanyi F, Ricol D, Thierry J;
XX
XX WPI; 2001-016103/02.
XX
DR
XX
XX Diagnosing carcinoma e.g. bladder or cervix carcinomas in a biological
XX sample such as tissue, bone marrow or body fluid, preferably from
XX animal or human, by identifying fibroblast growth factor receptor 3
XX mutations
XX
PS Disclosure; Fig 2q; 41pp; English.
XX
XX The identification of fibroblast growth factor receptor 3 (FGFR3)
XX mutations in a biological sample such as tissue, bone marrow or
XX body fluid e.g. urine, from a warm-blooded animal, preferably human
XX is useful for diagnosing carcinomas such as human bladder and cervix
XX carcinomas, or cancers associated with lung, breast, colon and skin.
XX The pharmaceutical preparations comprising agents which inhibit the
XX synthesis and expression of FGFR3 and so have an anti-proliferation
XX effect on carcinomas can be used to treat cancer.
XX
```

SQ Sequence 2427 BP; 428 A; 785 C; 820 G; 394 T; 0 other;

Query Match 3.7%; Score 118.6; DB 22; Length 2427;  
Best Local Similarity 56.1%; Pred. No. 4.3e-16;  
Matches 270; Conservative 0; Mismatches 199; Indels 12; Gaps 2;

```

OY 421 ccgcgttcacacagcccccacagatgagcgccggtgatcgacgcccgtggtgacg 480
    || || |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 451 ccttaactgacacagcgcccgagcgatgagacaagaactctgctgcgtgcgcgcgaac 510
OY 481 tccgtgcgtcacaagtgcgtgacgacgagcgacccctcgcccgacacacacgtgagaa 540
    |||| || || |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 511 accgtcgcgttcgcgtccgaagcgctggaaccacacccatccctcctcgtgctgaag 570
OY 541 gacgacacagcgtt-----gacgcgcgcacagagccgcgtgagcccgaggaagaag 591
    |||| || || |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 571 aacggaagggaggttccgcgcgcagccgcgcatggaagatcagctgcgcgtcagcag 630
OY 592 tggacactgagccctgagaacacctgcgcgcgagagacagcgcaaatatacactgcgcgtg 651
    |||| || || |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 631 tggagcctgtgtcatgaaagcgtgtgtccctcgacccgcgcacacacactgcgtcgtg 690
OY 652 tcgaacgcgcgcgcgcgcacacacacacacacacacacacacacacacacacacac 711
    |||| || || |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 691 gagaacaagtttgcgcagcatccgcgcagacgtacacgctggaacgtgtgagcgctcccg 750
OY 712 tccaagccgctgtctcacaagcagcaccctgcgaacacacacacacacacacacacac 771
    |||| || || |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 751 caacgcgcacatccctcagcgagcggtgcgcgcgcacacacacacacacacacacacac 810
OY 772 acgtccctcagtgcaaggtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 831
    |||| || || |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 811 gtggaattcacaactgcgaaggtgtacagtgacgcacacccacacacacacacacacac 870
DB 832 gtgga---gtacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 888
    |||| || || |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 871 gtggaagtgaaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 930
OY 889 t 889
DB 931 t 931

RESULT 35
AAA54423
ID AAA54423 standard; DNA; 2427 BP.
XX
AC AAA54423;
XX
AC 11-APR-2001 (first entry)
XX
DE Mutant (G377C) fibroblast growth factor receptor 3 (FGFR3-IIIB).
XX
KW Fibroblast growth factor 3 receptor; FGFR3; mutant; detection;
KW cancer; carcinoma; lung cancer; breast cancer; colon cancer;
KW skin cancer; bladder; cervix; human; ds.
XX
OS Homo sapiens.
XX
PN WO200068424-A2.
XX
PD 16-NOV-2000.
XX
PF 04-MAY-2000; 2000MO-EP04591.
XX
PR 05-MAY-1999; 99US-0132705.
XX
PA (CURT-) INST CURIE.
XX
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI Capellen D, Chopin D, Radvanyi F, Ricoul D, Thierly J;
XX

```

DR WPI: 2001-016103/02.

XX Diagnosing carcinoma e.g. bladder or cervix carcinomas in a biological  
PT sample such as tissue, bone marrow or body fluid, preferably from  
PT animal or human, by identifying fibroblast growth factor receptor 3  
PT mutations  
XX  
PS  
XX

Disclosure: Fig 2r; 41pp; English.

CC The identification of fibroblast growth factor receptor 3 (FGFR3)  
CC mutations in a biological sample such as tissue, bone marrow or  
CC body fluid e.g. urine, from a warm-blooded animal, preferably human  
CC is useful for diagnosing carcinomas such as human bladder and cervix  
CC carcinomas, or cancers associated with lung, breast, colon and skin.  
CC The pharmaceutical preparations comprising agents which inhibit the  
CC synthesis and expression of FGFR3 and so have an anti-proliferation  
CC effect on carcinomas can be used to treat cancer.

SQ Sequence 2427 BP; 428 A; 784 C; 820 G; 395 T; 0 other;

Query Match 3.7%; Score 118.6; DB 22; Length 2427;  
Best Local Similarity 56.1%; Pred. No. 4.3e-16;  
Matches 270; Conservative 0; Mismatches 199; Indels 12; Gaps 2;

```

OY 421 ccgcgttcacacagcccccacagatgagcgccggtgatcgacgcccgtggtgacg 480
    || || |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 451 ccttaactgacacagcgcccgagcgatgagacaagaactctgctgcgtgcgcgcgaac 510
OY 481 tccgtgcgtcacaagtgcgtgacgacgagcgacccctcgcccgacacacacacgtgagaa 540
    |||| || || |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 511 accgtcgcgttcgcgtccgaagcgctggaaccacacccatccctcctcgtgctgaag 570
OY 541 gacgacacagcgtt-----gacgcgcgcacagagccgcgtgagcccgaggaagaag 591
    |||| || || |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 571 aacggaagggaggttccgcgcgcagccgcgcatggaagatcagctgcgcgtcagcag 630
OY 592 tggacactgagccctgagaacacctgcgcgcgagagacagcgcaaatatacactgcgcgtg 651
    |||| || || |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 631 tggagcctgtgtcatgaaagcgtgtgtccctcgacccgcgcacacacacacacacacac 690
OY 652 tcgaacgcgcgcgcgcgcacacacacacacacacacacacacacacacacacacac 711
    |||| || || |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 691 gagaacaagtttgcgcagcatccgcgcagacgtacacgctggaacgtgtgagcgctcccg 750
OY 712 tccaagccgctgtctcacaagcagcaccctgcgaacacacacacacacacacacacac 771
    |||| || || |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 751 caacgcgcacatccctcagcgagcggtgcgcgcgcacacacacacacacacacacacac 810
OY 772 acgtccctcagtgcaaggtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 831
    |||| || || |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 811 gtggaattcacaactgcgaaggtgtacagtgacgcacacccacacacacacacacacac 870
OY 832 gtgga---gtacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 888
    |||| || || |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 871 gtggaagtgaaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 930
OY 889 t 889
DB 931 t 931

```

```

RESULT 36
AAA54424
ID AAA54424 standard; DNA; 2427 BP.
XX
AC AAA54424;
XX
AC 11-APR-2001 (first entry)
XX
DE Mutant (A393E) fibroblast growth factor receptor 3 (FGFR3-IIIB).
XX
KW Fibroblast growth factor 3 receptor; FGFR3; mutant; detection;
XX

```



Db	871	gtgagagtggaacgcgacgaagagtggtgcccgcgacggcacaccctacgttacgtgtcctaag	930
QY	889 t	889	
Db	931 t	931	
RESULT 37			
AAAS4407			
ID	AAAS4407	standard; DNA: 2427 BP.	
XX	AAAS4407;		
AC			
XX	11-APR-2001	(first entry)	
DT			
XX	Mutant (R248C) fibroblast growth factor receptor 3 (FGFR3-IIIB).		
DE			
XX	Fibroblast growth factor 3 receptor; FGFR3; mutant; detection;		
KW	cancer; carcinoma; lung cancer; breast cancer; colon cancer;		
KW	skin cancer; bladder; cervix; human; ds.		
XX			
OS	Homo sapiens.		
XX			
XX	WO200068424-A2.		
PN			
PD	16-NOV-2000.		
PF	04-MAY-2000; 2000WO-EP04591.		
PR	05-MAY-1999; 99US-0132705.		
XX			
XX	(CURI-) INST CURIE.		
PA	(CNRS ) CNRS CENT NAT RECH SCI.		
XX			
PI	Cappellen D, Chopin D, Radvanyi F, Ricol D, Thierry J;		
DR	WPI: 2001-016103/02.		
PT	Diagnosing carcinoma e.g. bladder or cervix carcinomas in a biological		
PT	sample such as tissue, bone marrow or body fluid, preferably from		
PT	animal or human, by identifying fibroblast growth factor receptor 3		
PT	mutations		
XX			
PS	Disclosure; Fig 2b; 41pp; English.		
XX			
CC	The identification of fibroblast growth factor receptor 3 (FGFR3)		
CC	mutations in a biological sample such as tissue, bone marrow or		
CC	body fluid e.g. urine, from a warm-blooded animal, preferably human		
CC	is useful for diagnosing carcinomas such as human bladder and cervix		
CC	carcinomas, or cancers associated with lung, breast, colon and skin.		
CC	The pharmaceutical preparations comprising agents which inhibit the		
CC	synthesis and expression of FGFR3 and so have an anti-proliferation		
CC	effect on carcinomas can be used to treat cancer.		
XX			
XX			
Sequence	2427 BP; 428 A; 783 C; 821 G; 395 T; 0 other;		
QY	Query Match	3.7%;	Score 117; DB 22; Length 2427;
Db	Best Local Similarity	55.9%;	Pred. No. 9.4e-16;
Matches	269; Conservative	0; Mismatches	200; Indels 12; Gaps 2;
QY	421	ccgcgttcacacacccctccaaagttagggcgccggtgatgcacgcccgtggtagc	480
Db	451	ccttacttgacacgcgcgcgagcgagtgacaaagaagctgctggcgtgcccgcgcacac	510
QY	481	tccgtgcggtcaaatgctgctgacgacggggaacccctgcggccgcagatcaagtgtgaag	540
Db	511	accgcgcgcttcggtgcccagccgctgacaaaccacatccctcatctcctgtgtag	570
QY	541	gacgacccggcctt-----gacggcccaagagcgctgtgcccagaagaagaag	591
Db	571	aacggcaggaggttcgcgcgcgcgaacccgcatcttgagagcatcaagctgcggcatcagcag	630



XX MO200046380-A2.  
 PN  
 XX 10-AUG-2000.  
 PD  
 XX 07-FEB-2000; 2000MO-US03166.  
 PE  
 XX 08-FEB-1999; 99US-0119002.  
 PR  
 XX (CHIR ) CHIRON CORP.  
 PA  
 XX Kavanaugh WM, Ballinger M;  
 PI WPI: 2000-514961/46.  
 DR P-PSDB: AAY97170.  
 DR  
 XX  
 PT New polypeptide comprising a fibroblast growth factor receptor  
 PT extracellular domain fused to a heterologous oligomerization domain for  
 PT treating FGF-, angiogenesis-, or FGF receptor-mediated disorders  
 PS  
 XX  
 PS Claim 16; Page 49-51; 70pp; English.

CC Novel fusion protein constructs comprise a fibroblast growth factor (FGF)  
 CC receptor (FGF-R) extracellular domain (ECD) lacking the immunoglobulin  
 CC (Ig) I segment fused to a heterologous oligomerization domain that  
 CC comprises an immunoglobulin Fe region, hinge region, CH1, CH2, CH3 or CH4  
 CC region, or light chain of an immunoglobulin molecule, or a peptide with a  
 CC leucine zipper motif. The Ig I segment is not necessary for binding of  
 CC acidic FGF and basic FGF (bFGF). The Ig I deletion further increases the  
 CC affinity for aFGF and heparin, protects the core of the molecule from  
 CC proteolysis, and abrogates the heparin requirement for aFGF binding. The  
 CC new fusion polypeptides are better FGF inhibitors than FGF-R monomer  
 CC proteins. The FGF-R-Ig Fc fusion dimers are active as FGF antagonists at  
 CC subnanomolar concentrations and were 20-fold more potent than the FGF-R  
 CC monomer protein as competitors of bFGF binding to immobilized FGF-Rs. The  
 CC fusion constructs are useful to treat FGF-, angiogenesis-, or  
 CC FGF-R-mediated disorders, such as tumorigenesis (e.g. bladder, breast,  
 CC lung, rectal, testis and cervical tumours), neovascularization (e.g.  
 CC diabetic retinopathy, neovascular glaucoma, wound healing and corneal  
 CC scarring) and hyper-proliferation of vascular smooth muscle cells (e.g.  
 CC postangioplasty and postatherectomy restenosis).

CC  
 XX  
 SO Sequence 1869 BP; 452 A; 564 C; 513 G; 340 T; 0 other;

Query Match 3.7%; Score 116.8; DB 21; Length 1869;

Best Local Similarity 51.2%; Pred. No. 9.7e-16;

Matches 333; Conservative 0; Mismatches 302; Indels 15; Gaps 2;

QY 203 ggaagcgttcgcgtcgtccgagggcgtgaaggtgaagcaggtgagcggaagatg 262  
 DB 227 gcaaccgaccgcgtacacagggaggtggaagtgcagagaccgtcccgagatc 286  
 QY 263 ccggcggtgtagctgtgcaaggccaccaggtctgagcagctgagcctacacacc 322  
 DB 287 ccggcctcctagctgtgcaaccagcagccctccggaagtgacacccactctccg 346  
 QY 323 tctgtctgtgtagtgcactagccaggaagagagcctggggccacagcctctctg 382  
 DB 347 tcaatgttcagatgtctccctctcgtcggagatgatgatgatgtactctctt 406  
 QY 383 ggggtcaagagagaccg-----ccagccagcagtgaggcagcagcgtctcacagc 436  
 DB 407 cagagagaaagaaagacagataacacaaacacccgtagctccatattgacatccc 466  
 QY 437 cctccaagaatagggcggtgtagcagagcggtgtgtagctccgtgagcctcagt 496  
 DB 467 cagaaagaatgaaagaatgtgcatgacgtgcgtgcgcagagcagtgaaagtcaat 526  
 QY 497 ggtgtgagcagcgagcccgccgagcagcagcagcagcagcagcagcagccttga 556  
 DB 527 gcccttcagtgaggaccccaaacccacacgtcgtgtgtgnaaatgtgcaaatctca 586

QY 557 cgc-----ggccagagcgctgagcccaagaagaagtgaagcactgagccctga 607  
 DB 587 aacctgaccacagattggaggtcactaacaggtccgtttagccacttgagcattcata 646  
 QY 608 agaacttgccgagcagagagcagcgcaaatlacacttgcgcgtgtgcagaccgcgagcg 667  
 DB 647 actctgtgtgccccctctgacaaggcaactacacttgatgttgagagatgagtacgca 706  
 QY 668 ccataacgcacactacaaggtgtagtgcagcagcgagcccgcttcacagccgtgtc 727  
 DB 707 gcatcaaccacacataccagctgtgtagtgcgtgagcggttcccttcacgcgcacatcgc 766  
 QY 728 cagcagcagcaccggtgacagcagcggtgagcttcgggggagcaacgcttcacgtca 787  
 DB 767 aagcaggtgtgcgcgcacacaaacagcagcgtgggtgagcaacgtgagttcatgtgta 826  
 QY 788 aggtgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 837  
 DB 827 aggtgtacagtgagccgcagcagcagcagcagcagcagcagcagcagcagcagcag 876

#### RESULT 40

AAx87179  
 ID AAX87179 standard; cDNA: 3343 BP.

AC AAX87179;

DT 27-SEP-1999 (first entry)

DE Fibroblast growth factor receptor 1 cDNA.

KW ZNF198-FGFR1, fibroblast growth factor receptor 1; oncogene;

KW Lymphoma; Leukaemia; SCLL; translocation; stem cell; diagnosis;

KW therapy; ss.

OS Homo sapiens.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC

Disclosure: Page 69-70; 76pp; English.  
 This is the nucleotide sequence of human fibroblast growth factor receptor 1 (FGFR1) cDNA. The FGFR1 gene is located on chromosome 8. A newly identified oncogene, ZNF198-FGFR1 (see AAX87177), incorporates a FGFR1 tyrosine kinase domain fused through t(8;13) translocation to ZNF198, a novel zinc finger gene of chromosome 13, and is associated with stem cell leukaemia/lymphoma syndrome (SCLL). Molecular characterisation of the oncogene and the encoded oncoprotein (see AAY06456) provides the ability to identify patients with t(8;13)(p11;q11-12) syndrome prior to onset of myeloid





```

XX OS Homo sapiens.
XX Key Location/Qualifiers
FH exon 210..467
FT /tag= a
FT /label= alpha
FT /note= "antisense oligomers pref. hybridise to this
FT region"
FT misc_difference 57
FT /tag= b
FT /note= "Identity of nucleotide is provisional"
FT misc_difference 117
FT /tag= c
FT /note= "Identity of nucleotide is provisional"
FT /note= "Identity of nucleotide is provisional"
PN MO9621471-A1.
PD 18-JUL-1996.
XX
XX 11-JAN-1996; 96WO-US00331.
XX
XX 10-JAN-1995; 95US-0371001.
XX
XX (GENT-) GENTA INC.
XX (MORR/) MORRISON R S.
XX
XX Brown BD, Morrison RS, Tseng BY:
XX
XX WPI: 1996-342063/34.
XX
XX Inhibiting growth of tumour cells - using an anti:sense oligomer for
XX the human fibroblast growth factor receptor-1 gene
XX
XX Disclosure: Page 54-55; 71pp; English.
XX
XX Antisense oligomers which hybridise to the human fibroblast growth
XX factor receptor 1 pre-mRNA and inhibit FGFRI expression are claimed.
XX The oligomers pref. hybridise to the alpha-exon region and are
XX useful for preventing growth of tumours, esp. glioma or glioblastoma
XX cells. The present sequence is that of the FGFRI gene.
XX
XX Sequence 2733 BP; 623 A; 811 C; 765 G; 532 T; 2 other:

```

```

Query Match 3.5%; Score 110.8; DB 17; Length 2733;
Best Local Similarity 50.8%; Pred. No. 2e-14;
Matches 333; Conservative 0; Mismatches 302; Indels 21; Gaps 2;

QY 203 ggaagcgccttcgcgtgtgcgcgaagggtgaagtgagcagtgagcggagagtg 262
DB 1111111111111111111111111111111111111111111111111111111
263 ccggcgtgacgtgtgcgcgaagggtgaagtgagcagtgagcggagagtg 322
DB 1111111111111111111111111111111111111111111111111111111
404 ccggcgtcctatgtcgtgcgaagggtgaagtgagcagtgagcggagagtg 463
QY 323 tcgtcgtgtgtgacattagccaggaagggtgaagtgagcagtgagcggagagtg 382
DB 1111111111111111111111111111111111111111111111111111111
464 tcaatgttctcagatgtctccctcctcgtgagtgatgatgatgatgatgatgat 523
QY 383 ggggtcagaagagaccgcg-----ccagccagcagtggtgagcagcagcgcctca 430
DB 1111111111111111111111111111111111111111111111111111111
524 cagaaggaagaagaacagtaaacacaaacacacagtaatgagcgtatgcatatgtga 583
QY 431 cacaagccctccacaagatgagcgcggtgacgcagcgccctgtgtgtgtgtgtgtgt 490
DB 1111111111111111111111111111111111111111111111111111111
584 catccccaagaagaagatggaagaagaattgcatgcatgcatgcatgcatgcatgcat 643
QY 491 tcaagtgcgtgtgcgcagcgagccctcgcgcagcagcagcagcagcagcagcagcag 550
DB 1111111111111111111111111111111111111111111111111111111
644 tcaatgtccttcctcagtgtagccccaacccaacacactgctgctgtgtgtaaatgtgcaag 703

```

```

QY 551 ccttgacgc-----gcccaagagccgctgtgagcccaaggaagaagtgtgacattga 601
DB 1111111111111111111111111111111111111111111111111111111
704 aattcaacctgaccacagaatttgaggtctacaaggtccgttatgtccacctgtgaacatca 763
QY 602 gccctgaagaacctgtgcgcgcgaggaagcagcagcgaataacattgcgcgtgtgtcgaaccgcg 661
DB 1111111111111111111111111111111111111111111111111111111
764 taatgacctctgtgtgtccctctgtgaagaagtgcaactcaaccctgcattgtggaagaatgat 823
QY 662 cgggcgcacatacagccacctacaaggtgtgattgtatccagcggaccgttccaaagcccg 721
DB 1111111111111111111111111111111111111111111111111111111
824 acggcagcatataccacacataccagcttgatgtcgtgtgagcgtgtccctccacgcgcga 883
QY 722 tgcctacagcagccaccccggtgaacacagcgtgtgacttgggggtgaccacgtctcttc 781
DB 1111111111111111111111111111111111111111111111111111111
884 tctgtcaagcaggggtgtgtccgcgaacaaacagtggtccctgtgtgcaacgttgagattca 943
QY 782 agtgcagagtggtgcgcagcagcagtgatcagtgatcagtgatcagtgatcagtgatcag 837
DB 1111111111111111111111111111111111111111111111111111111
944 tgttaaggtgtacagtgtagcccgacgcgcagcagcagcagcagcagcagcagcagcagcag 999

```

```

RESULT 44
AA027658
ID AA027658 standard; cDNA; 3525 BP.
XX
XX AA027658;
AC
XX
XX 03-FEB-1993 (first entry)
DT
XX
XX N-sam cDNA.
DE
XX
XX Tumour gene; bFGF receptor; antibody; detection; drug; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 271..2739
FT /tag= a
FT
XX
XX JP04190792-A.
XX
XX 09-JUL-1992.
PD
XX
XX 22-NOV-1990; 90JP-0316100.
PF
XX
XX 22-NOV-1990; 90JP-0316100.
PR
XX
XX (KOKU-) KOKURITSU GAN CENT SOCHO.
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
XX WPI: 1992-280111/34.
DR P-PSDB; AAR26337.
XX
XX Gene prod of N-sam tumour - and its recombinant-contg. nucleotide
XX sequence
XX
XX Claim 2; Fig 1; 19pp; Japanese.
XX
XX The sequence given is the N-sam tumour gene. The protein encoded by
XX this gene could be the receptor of human bFGF. The N-sam protein
XX and peptide fragments of N-sam are useful as anti-tumour drugs.
XX Anti-N-sam peptide antibodies are useful for the detection of the N-
XX sam gene product, in the purification of the N-sam gene product and
XX in the diagnosis of tumours and anti-tumour drugs.
XX
XX Sequence 3525 BP; 781 A; 1046 C; 1002 G; 696 T; 0 other:

```

```

Query Match 3.5%; Score 110.8; DB 13; Length 3525;
Best Local Similarity 50.8%; Pred. No. 2.1e-14;
Matches 333; Conservative 0; Mismatches 302; Indels 21; Gaps 2;

QY 203 ggaagcgccttcgcgtgtgcgcgaagggtgaagtgagcagtgagcggagagtgatg 262

```

```

Db      497 gcaaccgaccgcatcacagggagaggtgaggtgcaagactccgtgcccagact 556
QY      263 ccgagcggtacgtgtgtaagggcccaacggtcgcagcgtgagcgtcaatacacc 322
Db      557 ccgagcctctatgtctgttaacagcagccctcgcgagtgacacacactactctccg 616
QY      323 tcgtcgtcgtgatacatagccagggagagagccttggtggccgacagcctcctg 382
Db      617 tcaatgtttcaatgactctccctccctcgcgagatgatgatgatgatgatgatgatgat 676
QY      383 gggtgtcaagagagaccgcg-----ccagccagcagtggtggccagcagcgtctca 430
Db      677 cagagggagaagaacagataacacaaacacgltatgccgtgtgtccatatttggga 736
QY      431 cacagccctcaagatgtaggcgcggtgtatgcagcagccgtgtgtagctcgtgcgac 490
Db      737 catcccaagaagaatgtgaaagaatgtcatgtcagtgccgctgcacaagacgtgaagt 796
QY      491 tcaagtgtcgtgcagcagggcagccctcgcgcgcacatcaagtgtagtgaagacagcagg 550
Db      797 tcaaatgtcccttcagtcggagcccaacccacacgtcgctgtgtgaaatgtgcaag 856
QY      551 ccttgacgc-----gcccaaggcgcgtgagcccaagaagaagaagtgtgacactga 601
Db      857 aattcaaacctgacacagaaattgtgaggtctaaaggtccgtgtatgccactgtgagcatca 916
QY      602 gcttgaagaacctgtgcgcgagagacagcgcaaatcacctgcgcggtgtcgaaccgcg 661
Db      917 taatgtgactgtgtgtgctcctctgacaagggcaactacacactgattgtgtgaatgagt 976
QY      662 cgggscgcatcaacgcacactacaaagtgtgataccagcagccggtcccaagcccg 721
Db      977 acggcagcatcaaacacacatcacagctgtgtgtgagcggtccctccacgcgcga 1036
QY      722 tgcctcagagcagcagccagcgtgacacagcagcagcgtgagcttgcgggggagccagctctcc 781
Db      1037 tccctcagagcaggggttgcgcgcacacaaacagtgccctgggtgagcaagtgagttca 1096
QY      782 agtgcagagtgtgcagcagcgtgaaagcgcgtgtatccagtgctgtgaagcgcgtgag 837
Db      1097 tgtgtgaagtgtatagcgtgaccgcgagcgcacatccagtgctgtgaagacatcgag 1152

```

## RESULT 45

AAQ21003 standard; cDNA; 2662 BP.

AAQ21003:

20-MAY-1992 (first entry)

flag receptor protein gene.

fibroblast growth factor receptor; heparin binding proteins;

tyrosine kinase; fms-like gene; CSF-1; PDGF; ss.

Homo sapiens.

Key Location/Qualifiers

FT CDS 87..2546

FT misc\_feature /tag= a

FT misc\_feature /product= flag

FT misc\_feature /tag= b

FT misc\_feature /note= "insertion; see comments"

FT misc\_feature /tag= c

FT misc\_feature /note= "deletion; see comments"

```

XX      03-JUL-1991; 91MO-US04745.
PE      06-JUL-1990; 90US-0549587.
PR      (RORE ) RORER INT HOLDINGS.
PA      Dloune CA, Crumley G, Juye MC, Schlessinger J;
PI      WPI; 1992-056827/07.
DR      P-PSDB; AAR20750.
XX      New fibroblast growth factor receptor proteins - useful in
PT      treating GF-mediated conditions e.g. angiogenesis of tumours,
PT      mitogenic effects in psoriasis, arthritis
XX      Claim 10; Fig 7; 65bp; English.
XX      The cDNA sequence was obt'd. using as template the partial flg known
CC      sequence (Ruta et al., Oncogene, 3: 9-15). A full length flg cDNA
CC      clone was obt'd. by repeated rounds of PCR using 8 primers (AAQ23610-
CC      17), and was cloned into pMJ30. NIH 3T3 cells were cotransfected
CC      with a 1:20 mixture of pSV2neo and an flg expression vector contg.
CC      the flg probe inserted into pMJ30 immediately downstream of the SV40
CC      promoter and cytomegalovirus enhancer. Clones of flg transfected cells
CC      have increased binding for arGF receptors, the best being clone N1926
CC      which was used for DNA sequencing. The sequence shown contains an
CC      apparent insertion (see features) which disrupts the reading frame.
CC      This is restored, however, by a deletion downstream in the sequence.
CC      No reference is made by the authors to these discrepancies, which are
CC      therefore assumed to be typing errors. Receptor proteins encoded by
CC      this cDNA sequence may be used in pharmaceutical compns. to inhibit
CC      undesirable heparin-binding growth factor mediated cellular responses
CC      or to inhibit the binding of an opportunistic pathogen to human cells.
CC      Such undesirable responses may be growth factor stimulated angiogene-
CC      sis and vascularisation of tumours, mitogenic effects in psoriasis,
CC      arthritis, atherosclerosis and benign prostatic hypertrophy. The
CC      derived flg protein may also be used for screening drugs for treatment
CC      of such problems.
CC      See also AAQ21004.
XX      Sequence 2662 BP; 613 A; 787 C; 740 G; 522 T; 0 other;
SQ

```

## Query Match

Best Local Similarity 50.6%; Score 109.2; DB 13; Length 2662;

Matches 332; Conservative 0; Mismatches 303; Indels 21; Gaps 2;

```

QY      203 ggaagccttcgcgtgtcgcgaggggtcgtgaaggtgaaagcaggtgtgagcggagatg 262
Db      314 gcaaccgaccgcatcacagggagaggtgaggtgcaagactccgtgcccagact 373
QY      263 ccgagcggtgactgtgtgcaagggccacacagcgtcgcgacgctgagtgcaatacacc 322
Db      374 ccgagcctctatgtctgttaacagcagccctcgcgagtgacacacactactctccg 433
QY      323 tcgtcgtcgtgatacatagccagggagagagccttggtggccgacagcctcctg 382
Db      434 tcaatgtttcaatgactctccctccctcgcgagatgatgatgatgatgatgatgatgat 493
QY      383 gggtgtcaagagagaccgcg-----ccagccagcagtggtggccagcagcgtctca 430
Db      494 cagagggagaagaacagataacacaaacacgltatgccgtgtgtccatatttggga 553
QY      431 cacagccctcaagatgtaggcgcggtgtatgcagcagccgtgtgtagctcgtgcgac 490
Db      554 catcccaagaagaatgtgaaagaatgtcatgtcagtgccgctgcacaagacagtgagt 613
QY      491 tcaagtgtcgtgcagcagggcagccctcgcgcgcacatcaagtgtagtgaagacagcagg 550
Db      614 tcaaatgtcccttcagtcggagcccaacccacacactgcgtgtgttgaataatgtgcaag 673
QY      551 ccttgacgc-----gcccaaggcgcgtgagcccaagaagaagaatgtgacactga 601

```

```
Db 674 aattcaaacctgacccaagaatttgaggtacaaaggtccgttatgtccacctggagcatca 733
Qy 602 gccctgaagaacctgcygcccggaggaacgcgcaaatatacaactgcgcggtgtcgaaccgcg 661
Db 734 taatgacctgtgtgtgccctctgacaagggcaactacacctgcaltgtgagaatgagt 793
Qy 662 cgggagccatcaacgccaactacaaggttgatgtatccagcggaaaccgltccaagccg 721
Db 794 acggcagcatcaaccaacataccagcttgatgtcgttgagcgcgtccctcaaccgccca 853
Qy 722 tgcacaaagcagcagcccgctgaacacgacgctggaactcggggggaacacgtccctcc 781
Db 854 tccctgaagaagcaggtgtgcccgcacaacaacagtggtccctggttagcaacgtggagttca 913
Qy 782 agtgcagaagtgccgacgacgagcgtgaagccggtgattccagtgctgaagcgcgtggag 837
Db 914 tgtgtaaggtgtacagtgaccgcgacgcgcacatccagtgctaaagcacatcgag 969
```

Search completed: August 17, 2001, 04:21:49  
Job time: 6980 sec